TECHNISCHE UNIVERSITÄT
CHEMNITZ

# Characterizing predictive auditory processing with EEG

von der Fakultät für Naturwissenschaften der Technischen Universität Chemnitz
genehmigte Dissertation zur Erlangung des akademischen Grades

doctor rerum naturalium

(Dr. rer. nat.)

vorgelegt von M.Sc. Martin Reiche
geboren am 15.03.1987 in Gotha
eingereicht am 21.12.2016

Gutachter:
Prof. Dr. rer. nat. habil. Alexandra Bendixen
Prof. Dr. phil. habil. Erich Schröger

Tag der Verteidigung: 09.06.2017
Veröffentlicht unter: http://nbn-resolving.de/urn:nbn:de:bsz:ch1-qucosa-226275

# Bibliographische Beschreibung

Reiche, Martin

**Titel:** Characterizing predictive auditory processing with EEG

Dissertation an der Fakultät für Naturwissenschaften der Technischen Universität Chemnitz, Institut für Physik, Dissertation, 2016

118 Seiten; 30 Abbildungen; 3 Tabellen; 163 Literaturzitate

## Referat

Predictive coding theorizes the capacity of neural structures to form predictions about forthcoming sensory events based on previous sensory input. This concept increasingly gains attention within experimental psychology and cognitive neuroscience. In auditory research, predictive coding has become a useful model that elegantly explains different aspects of auditory sensory processing and auditory perception. Many of these aspects are backed up by experimental evidence. However, certain fundamental features of predictive auditory processing have not been addressed so far by experimental investigations, like correlates of neural predictions that show up before the onset of an expected event. Four experiments were designed to investigate the proposed mechanism under more realistic conditions as compared to previous studies by manipulating different aspects of predictive (un)certainty, thereby examining the ecological validity of predictive processing in audition. Moreover, predictive certainty was manipulated gradually across five conditions from unpredictable to fully predictable in linearly increasing steps which drastically decreases the risk of discovering incidental findings. The results obtained from the conducted experiments partly confirm the results from previous studies by demonstrating effects of predictive certainty on ERPs in response to omissions of potentially predictable stimuli. Furthermore, results partly suggest that the auditory system actively engages in stimulus predictions in a literal sense as evidenced by gradual modulations of pre-stimulus ERPs associated with different degrees of predictive certainty. However, the current results remain inconsistent because the observed effects were relatively small and could not consistently be replicated in all follow-up experiments. The observed effects could be regained after accumulating the data across all experiments in order to increase statistical power. However, certain questions remain unanswered regarding a valid interpretation of the results in terms of predictive coding. Based on the current state of results, recommendations for future investigations are provided at the end of the current thesis in order to improve certain methodological aspects of investigating predictive coding in audition, including considerations on the design of experiments, possible suitable measures to investigate predictive coding in audition, recommendations for data acquisition and data analysis as well as recommendations for publication of results.

**Schlagworte:** predictive coding, sensory processing, auditory perception, ecological validity, gradual manipulation, EEG, event-related potentials, passive listening, human subjects

# Acknowledgements

First and foremost I would like to thank my supervisor Alexandra Bendixen for entrusting me to work on the current thesis, for granting me unique insights into her scientific work and for her constant support throughout the last four years. I would also like to thank Erich Schröger for assessing the current thesis and for introducing me to his scientific work which ultimately led me to further investigate the current topic. I respectfully thank Andreas Widmann for his methodological suggestions which substantially contributed to the development of the experimental designs and for providing valuable advice on several issues like data analysis or technical problems in times of need. Furthermore, I would like to thank Stefan Debener for granting me access to his laboratory where one of the experiments of the current thesis was conducted. I would also like to thank Bradley Jack for proofreading the current thesis. I gratefully acknowledge the German Research Foundation (DFG) and Chemnitz University of Technology for funding my project. Finally, I would like to thank all the people who eagerly supported me during the preparation of experimental sessions and with data acquisition. I especially thank Thomas Baumann, Christina Brauckmann, Sara Calado, Maria Stuckenberg, Christiane Neubert, Teresa Klohn, Markus Pelz and Carina Sander.

Chemnitz, December 21, 2016, Martin Reiche

# Contents

# 1 Introduction

Perception is the fascinating process that enables us to interact with our environment. We can see rich and colorful natural scenes and can distinguish uncountable different objects by the help of vision, yet we only see a tiny fraction of the electromagnetic spectrum. We can perceive waves of compressed and decompressed air which we interpret as sounds through hearing. We are able to hear sounds in a frequency range from about 20 Hz to 20 kHz and moreover, we can interpret complex compositions of a variety of different sounds as music. Maybe most strikingly, the ability of hearing enables us to communicate using complex combinations of sounds produced by the human vocal tract to exchange information with other individuals. Our vestibular system tells us whether we are moving or at rest. Muscle spindles provide us with information about the relative positions of our limbs through proprioception. Thermoreceptors in our skin enable us to distinguish between higher and lower levels of thermal energy. On top of all of that, we are able to sense a huge variety of different chemical compounds with the help of olfaction and gustation. These senses open up a world of different tastes and odours and not only determine how we perceive the food we eat but also influences how we emotionally experience the world around us. All these sources of information provide a very detailed picture about what we perceive as reality. Not only does each individual sensory system provide a lot of information but additionally, these systems influence each other. Perceptual information from one channel can easily interact with one or more other perceptual systems. Such processes are commonly known as multimodal or multisensory perception. One very intriguing example of multisensory perception is the McGurk effect (McGurk & MacDonald, 1976) in which visual information can interfere with auditory processing which can result in a blended percept or a fusion of the competing information. When you watch a video of a person uttering the syllables '*ba-ba*' but the video is dubbed with the audio of the person saying '*ga-ga*', the resulting percept will be a blended version of the conflicting information: you will actually hear the syllables '*da-da*'. Such interactions can exist between multiple different perceptual subsystems which should serve as an example of how complex perceptual processes can be. From comparatively simple mechanisms, like sensory transduction of periodic signals by hair cells in our inner ear to highly perplexing multisensory phenomena, like the McGurk effect, a lot of research has been conducted to gain an understanding of the underlying mechanisms of perception.

Thanks to the early work of researchers like Alan Hodgkin and Andrew Huxley (Hodgkin & Huxley, 1945), we now have a profound understanding of the basic units of perception. We know that perception is essentially carried out by networks of neurons within our central nervous system. These neurons can transduce information by changing states of their membrane potentials and when interconnected, they basically follow the principles of Boolean logic. Hence the fundamental building blocks of perception are very well known and the outcome of these comparatively simple mechanisms makes up everything we experience and determines how we perceive the world around us. However, it is to a large part still a subject of speculation how in detail such rich perceptual experiences can be formed by these comparatively simple mechanisms. Despite this explanatory gap, many theories have been proposed and useful models have been brought forward to gain a better understanding of the mechanisms underlying perceptual processes.

One of the first experimental approaches to systematically investigate perceptual systems was Psychophysics established by Gustav Theodor Fechner (exemplified by his influential monograph published in 1860). Psychophysics traditionally aimed at examining the relation between changing physical intensities of stimuli and the corresponding subjective intensity change within perception, which was considered as unquantifiable to this date (Müsseler, 2008). Nowadays Psychophysics is commonly used within the framework of threshold measurements, and as such has many applications in the context of research but also in engineering and the consumer industry. Modern applications of Psychophysics are expanded by concepts like ideal observer analysis (Tanner & Birdsall, 1958) and signal detection theory (Tanner & Swets, 1954; Swets, 1964). A common practical example of technology that is guided by Psychophysics is lossy compression within the field of digital signal processing. By exploiting the rules of Psychophysics, lossy compression enables to significantly reduce the amount of data used to store audio and video files without a decrease in subjectively perceived signal quality. Other approaches to explain perceptual processes like Gestalt Psychology aimed at describing the organizational principles for determining which of the sensory signals belong together (i.e., were emitted from the same object in the environment) and which signals belong to different objects (e.g. Wertheimer, 1923). Those principles are expressed in the Gestalt laws like the law of proximity or the law of similarity (Müsseler, 2008) which were often studied in visual perception, but in fact they can be applied to other modalities like auditory perception as well.

Approaches like Psychophysics and Gestalt Psychology helped to understand the relationship between physical stimuli and resulting perceptual phenomena to a great extent. They can formulate functional relationships between the physical and the perceptual world and predict how changes in the physical world might translate into changes in our perception. However, they solely rely on measuring behavioral parameters without the ability to assess what happens in between the physical world and the behavioral outcome. But how exactly is perception implemented in the brain and what are the operations the brain has to perform to get from sensory signals impinging upon our sensory epithelia to a coherent and informative percept? New technologies have been developed to breach this gap and to get further insights about the neural implementations of perception, like Functional magnetic resonance imaging (fMRI), Positron emission tomography (PET), Magnetoencephalography (MEG) and Electroencephalograhy (EEG). These techniques paved the way for a completely new field of research: cognitive neuroscience, which aims at finding neural correlates of cognitive and perceptual processes and sets out to explore the underlying neural mechanisms. The increase in computational performance throughout the last decades allows us furthermore to simulate complex models of proposed neural mechanisms. Simulations range from small neural networks with hundreds of units (Wacongne, Changeux, & Dehaene, 2012) to simulations of whole cortical columns consisting of tens of thousands of units with remarkably high anatomical fidelity (Markram, 2006). These advancements allow to go beyond what is experimentally accessible and to run multiple possible scenarios testing a variety of hypotheses and to inform theories about perceptual processing. One of the leading theories which addresses perceptual processing in the brain is predictive coding theory (Friston, 2005) which has dominated experimental research in the last few years. The current thesis aims at further investigating how predictive coding theory can serve as a model to explain auditory perceptual processing.

In this section of the thesis, first, a general introduction to predictive coding theory will be given. In the following part, basic principles of auditory perception will be introduced and prediction will be discussed as a plausible mechanism of auditory perception. Later on in this section, an overview on previous research and the current state of knowledge will be given regarding empirical correlates of prediction. Since all of the experimental work conducted within the framework of the current thesis was done by means of EEG, special emphasis will be put on electrophysiological research. On this basis, the main goals of the thesis will be presented and the overall logic of the conducted experiments will be explained subsequently. The experimental work of the current thesis will be presented in three major parts. In Section 2, traditional electrophysiological correlates of predictive auditory processing were investigated by expanding an established experimental paradigm introduced by Bendixen, Schröger, and Winkler (2009). Two experiments were designed to systematically investigate the influence of different aspects of predictability on perceptual processing of acoustic stimuli measured with event-related potentials (ERPs). In Section 3, an attempt was made to overcome certain theoretical and methodological shortcomings of previous research by investigating ERP-correlates of prediction within the pre-stimulus time range with the aim of finding contributions to ERPs that are 'purely' driven by prediction (as compared to post-stimulus ERPs which are inevitably contaminated by the processing of the stimulus itself). Two further experiments will be introduced in this section. The first aimed at investigating the temporal dynamics of the observed effects and the other experiment was designed to investigate the influence of rare tone omissions within the experimental stimulus sequences. Deviant events like this are commonly used to investigate perceptual processes within EEG-research (Joutsiniemi & Hari, 1989; Raij, McEvoy, Mäkelä, & Hari, 1997; Hughes et al., 2001; Todorovic, van Ede, Maris, & de Lange, 2011) and as such were also used in the current thesis. However, these rare events lead to subjectively perceived disruptions in the tone sequences that were used in the experiments. This last experiment was conducted to rule out that unexpected effects were introduced by omissions of experimental stimuli in the paradigms applied here.

There has been a growing awareness of insufficient reproducibility and reliability within biomedical (Ioannidis, 2005) and psychological (Anderson et al., 2015) research in the last few years. According to a review by Barch and Yarkoni (2013), there are several reasons for this so-called 'replication crisis', like conflicts of interest (Bakker & Wicherts, 2011), misaligned incentives and questionable research practices (John, Loewenstein, & Prelec, 2012) which results in what is often referred to as 'p-hacking' (Simmons, Nelson, & Simonsohn, 2011) and ubiquitous low power (Button et al., 2013). According to Barch and Yarkoni (2013), such problems may lead to a higher incidence of false positive results which in turn is a major contribution to publication bias. To take measures against aforementioned problems, like insufficient power, in Section 4 statistical analyses were carried out on accumulated data across Experiments wherever this was applicable. A comprehensive discussion of the overall results, interpretations and suggestions for further research will be given in Section 5.

## 1.1 An introduction to predictive coding theory

For a long time the prevailing notion about the fundamental principles of perception was characterized by models of sequential computational operations. It was commonly assumed that perception is carried out in a retrospective manner (Bregman, 1990): physical stimuli are detected at the sensory organs and transduced into a code of neural signals. On early levels of perception, certain basic features are extracted, like tone pitch in audition or the orientation of edges in vision. In later perceptual processes, these features are bound together and get categorized on higher, progressively abstract levels. Finally, the resulting percepts are coregistered with memory information. According to such a theory of perception, the brain continuously analyzes its sensory input by carrying out those steps one after another. Strictly speaking, this implies that at every point in time, each stimulus gets processed the same way even though the stimulus does not change. The system would perform these steps over and over again even though there would be no gain of information. There are however several objections to this notion. From a metabolic perspective, neural computations are very expensive. The human brain causes nearly a quarter of the metabolic costs of the whole organism (Leonard, Robertson, Snodgrass, & Kuzawa, 2003). Throughout ontogeny, several mechanisms have evolved to optimize the efficiency of the brain (i.e. to increase computational performance and at the same time keep metabolic costs as low as possible), like neural pruning (Gazzaniga, Ivry, & Mangun, 2009). Current research indicates that the brain works in a highly efficient manner on a metabolic level (as compared to electrical circuits) through the advantageous properties of balanced networks with inhibitory and excitatory activation (Sengupta & Stemmler, 2014).

Why should a system that is highly specialized to perform complex neural computation in nearly real-time, which at the same time tries to be metabolically as cost-efficient as possible, rely on such an inflexible and lavish modus operandi? In other words, why should the brain process the same stimulus over and over again instead of using the information that is already processed and save resources to process only new information? Other theoretical approaches to explain perceptual processing have been developed which take questions like these into account. An alternative theory which has recently gained more and more interest proposes that the perceptual system actively makes predictions about future sensory events based on models which were shaped by preceding input (e.g. Friston, 2005; Friston & Kiebel, 2009; Bar, 2007). Being able to make predictions about events which are likely to happen in the future allows the system to save resources because the expected events do not have to be processed entirely due to the fact that certain features of the upcoming event are already known (e.g. Sinkkonen, 1999). We continuously make predictions in order to successfully reach our goals, we predict the behavior of others in order to adequately prepare for our reactions and we make predictions about physical objects in the world surrounding us in order to adapt to our environment. Recent investigations paid special attention to predictive processes not only within higher cognitive operations but also within more basal and automated mechanisms of the nervous system, like the processing steps leading to conscious perception. This led to theoretical frameworks and mathematical models which aim at explaining predictive coding within perception.

One of the leading proposals was introduced by Karl Friston (2005). This theory is based on the concept of perceptual inference, dating back to Helmholtz (1867). Perceptual inference means to infer from sensory input to what could have most likely caused this input. Thus, the brain is constantly testing hypotheses about the physical causes of the sensory information at hand. The recognition of these causes is accomplished by internal models which are shaped by experience about the environment. Predictive coding theory relies on such inferential processes and assumes that perception is based on empirical Bayes models, predicting expected upcoming sensory input and comparing it with the actual input (Clark, 2013; Srinivasan, Laughlin, & Dubs, 1982; Rao & Ballard, 1999). This system is assumed to be based on a cortical hierarchy in which lower cortical structures have forward connections to higher structures delivering sensory input whereas higher structures can modulate lower parts through predictions using backward connections. The aim of the system is to ultimately minimize the prediction error by adjusting the model in order to optimize the predictions (Friston & Kiebel, 2009). According to Kanai, Komura, Shipp, and Friston (2015), there is anatomical and physiological evidence supporting predictive coding in the brain (Friston, 2008; Mumford, 1992). Perceptual systems might be based on canonical microcircuits and hierarchical predictive coding (Bastos et al., 2012) and there are also plausible models of predictive coding in the motor system (Adams, Shipp, & Friston, 2013). The functional principle of such models is based on neural representations on higher cortical levels generating predictions about sensory evoked activity in lower levels of the neural hierarchy. These neural top-down predictions are passed down the cortical hierarchy and are compared with the actual bottom-up sensory input. Parts of the bottom-up signal which could not be explained by the top-down prediction are passed on to higher cortical levels in the form of prediction error signals which will be used to refine the neural model by optimizing the predictions. The system aims at explaining as much of the sensory signal as possible and to minimize the prediction error on all levels of the neural hierarchy. These operations are assumed to be carried out by cortical pyramidal cells on several hierarchically organized layers (Kanai et al., 2015).

Predictive coding allows the brain to reduce the resources needed for processing of information to a minimum by enabling the system to use information that is already available (the prediction) and focus only on parts of the signal that could not be explained away by the neural predictions (the prediction error). Only new information needs to be processed and is used to update and adjust the generative model of causes which attempts to explain the sensory information. In comparison to retrospective processing, as explained at the beginning of this section, such a prospective mechanism of perception would be superior in many regards, like efficiency, metabolic energy consumption or processing speed. In the following section, an implementation of predictive coding will be discussed in greater detail within the framework of auditory perception. Further examples will be given to illustrate the advantages of predictive vs. retrospective processing in perception.

## 1.2 Predictive coding in audition

The concept of predictive mechanisms in perception becomes especially plausible in audition. This has to do with the physical nature of sound (see also Bendixen, SanMiguel, & Schröger, 2012). The physical cues which can be used by the auditory system to extract information are frequency, intensity, spectrotemporal information, like timbre and location which is coded by interaural time and intensity differences. All of these features carry information only when evaluating them over time. Due to the unfolding in time, as soon as acoustic information has reached our ears, it disappears again and is no longer available for “re-inspection”. The system has to integrate this information over a certain period of time and rapidly extract certain regularities (e.g Bendixen, Prinz, Horváth, Trujillo-Barreto, & Schröger, 2008; Bendixen & Schröger, 2008; Paavilainen, Jaramillo, Näätänen, & Winkler, 1999). This challenge becomes even more apparent if you consider that all the information from multiple sound sources available at one point in time is superimposed. As Bregman (1990) ingeniously shows in an analogy: it is as if you tried to tell what is happening on a huge lake, how many boats there are, how big they are, how far they are away, whether the wind is blowing and so on, just by looking at the waves at the lakeside. When visual objects are superimposed, you can obviously only see what is on top or in front with the exception of transparent objects like glass panes (as opposed to auditory objects which can in general be considered "transparent"). Furthermore, visual information is often available for a longer period of time so the visual system can reanalyze stimuli. This is simply not possible for the auditory system in most of the cases. Taking all this into account, it seems highly inappropriate to assume that the auditory system is just sequentially analyzing every single point in time and restoring the information without the use of any sophisticated mechanism. The current thesis is an attempt to further probe whether predictive processing could be such a mechanism and to give further insights about its characteristics.

One possible implementation of a neural network based on predictive coding within the auditory cortex is depicted in Figure 1.1. This model was brought forward by Wacongne et al. (2012). The authors used this implementation to provide a computational model composed of Izhikevich neurons (Izhikevich, 2003) which was able to explain the sources of a common electrophysiological correlate of predictive coding in auditory perception, namely the so-called mismatch negativity (MMN). The MMN, commonly associated with prediction error signals, was first discovered by Näätänen, Gaillard, and Mäntysalo (1978) and will be further discussed in the next section among other correlates of predictive auditory processing. In the model by Wacongne et al. (2012), subpopulations of neurons, sensitive to different tone frequencies, are hierarchically organized in columns representing tonotopically organized cortical columns in the primary auditory cortex. These columns receive excitatory input from thalamic neurons responding maximally to their sensitive frequencies. According to the authors, the thalamic signals enter the cortex through feedforward connections to granular pyramidal cells (lamina IV). Apart from the thalamic input (sensory information), these cells receive input through

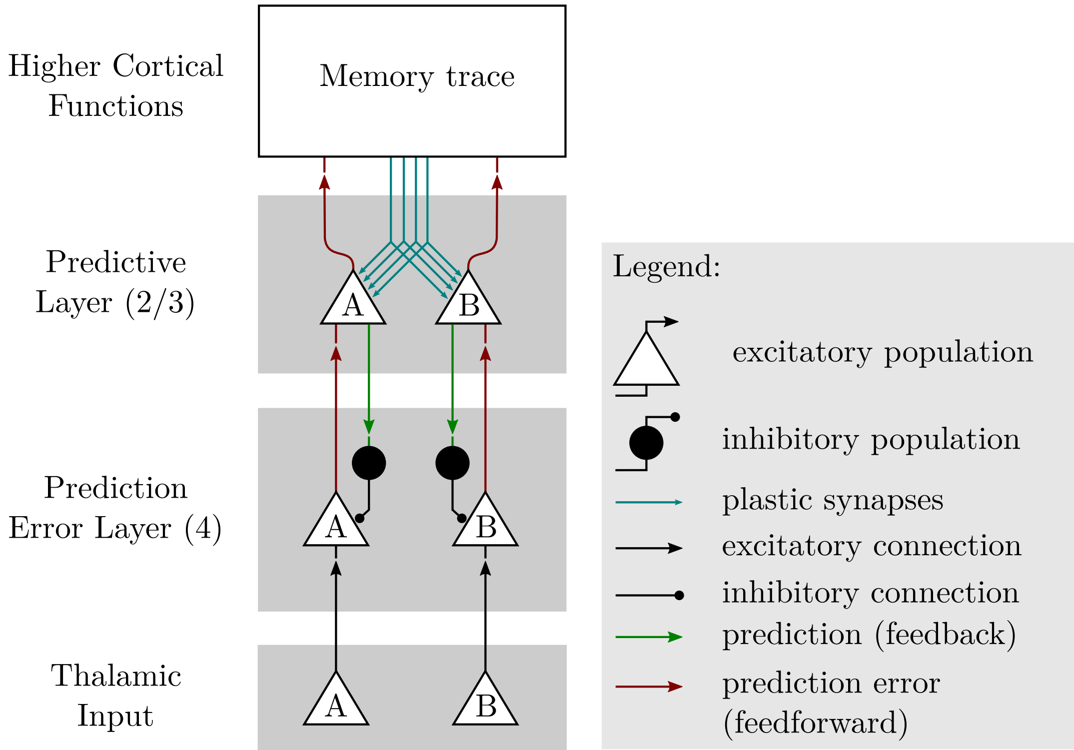


Figure 1.1: Model implementation of predictive coding in primary auditory cortex, based on Wacongne et al. (2012): hierarchically structured neuronal populations are ordered in columns, representing tonotopically organized cortical columns in primary auditory cortex (here: two columns maximally sensitive to tones of frequency A and B). Granular pyramidal cells in the prediction error layer receive sensory input from thalamic neurons (black arrows) and inhibitory predictive input from supragranular pyramidal cells in the predictive layer (green arrows). The difference between expected sensory input and actual thalamic input is fed forward in the form of prediction error activity (red arrows) which is sent to higher cortical levels in order to adjust the predictive model which is based on previous sensory input. The system aims at optimizing the predictions by minimizing the prediction error.

feedback connections from higher cortical layers (predictive information). These predictions are generated by excitatory populations of supragranular pyramidal cells located in laminae II/III. Therefore this layer will be referred to as predictive layer. Cells in this layer receive input from higher cortical layers associated with sensory memory functions. Here, the actual predictive model is assumed to be shaped based on regularities in the previous sensory input. If the model predicts a stimulus of a certain frequency, cells specific to this frequency within the predictive layer receive excitatory input from higher cortical layers. These will then form a prediction of the expected upcoming sensory activation which gets fed back to frequency specific cells in the granular layer. The predictive feedback activation gets sign-inverted by inhibitory interneurons. As a consequence, the prediction arrives at layer IV neurons as inhibitory activation. If sensory information that is fully predictable reaches these neurons through the excitatory thalamic inputs, it gets cancelled out by the inhibitory predictive activation from the predictive layer. If the sensory information could not or only partly be explained by the prediction, the difference of the expected input and the actual sensory input will be forwarded to the predictive layer by excitatory feedforward connections. This activation is commonly called the prediction error and is assumed to be generated by neurons in layer IV. Therefore this layer will be referred to as the prediction error layer. The prediction error activity is used on higher levels of the hierarchy to adjust the generative predictive model and to optimize the predictions in order to minimize the prediction errors.

With this model, Wacongne et al. (2012) were able to demonstrate that predictive coding in the primary auditory cortex can account for major empirical properties of the MMN, like frequency-dependent or duration-dependent response to rare deviants (Näätänen, Paavilainen, & Reinikainen, 1989), a response to unexpected repeats in alternating sequences (Horváth & Winkler, 2004), a lack of consideration of the global sequence context (Bekinschtein et al., 2009), a response to sound omission (e.g. Joutsiniemi & Hari, 1989; Yabe, Tervaniemi, Reinikainen, & Näätänen, 1997; Raij et al., 1997; Hughes et al., 2001) and a sensitivity of the MMN to N-methyl-D-aspartate (NMDA) receptor antagonists (Ehrlichman, Maxwell, Majumdar, & Siegel, 2008; Tikhonravov et al., 2008, 2010).

In the next section, an overview of empirical evidence of predictive auditory processing will be given. A strong focus will be set on electrophysiological results within EEG measurements with a special emphasis on ERP measures.

## 1.3 Electrophysiological correlates of predictive auditory processing

If there is predictive processing of auditory sensory signals, objective signs of this should be observable in a modulation of brain signals at or slightly before stimulus-onset. Such early and short-lived effects can hardly be investigated using imaging techniques, like fMRI due to the relatively imprecise temporal resolution of blood-oxygen-level dependent (BOLD) signals. On the contrary, electrophysiological measures which can be obtained with EEG or MEG provide a very good temporal resolution and therefore prove useful to investigate such early effects. A lot of research has been conducted using ERPs obtained by EEG measurements. ERPs can be used to investigate the processing of certain time-locked cognitive, sensory or motor events by applying a simple averaging technique (Luck, 2005). In the following, some of the observations investigating predictive processes by means of electrophysiological measures will be sketched out to further illustrate the motivation of the current thesis.

Furthermore, it should be clarified that predictive processing in perception is assumed to take place in an automatic manner. Perceptual processes do not explicitly require individuals to actively invest resources in order to obtain an internal model of the world surrounding them. That is for example, we do not need to employ attentional resources and we do not have to exert active cognitive control in order to successfully perceive the world around us (Bendixen et al., 2012). Quite the opposite seems to be the case: up until a certain point, basic perceptual processing seems to take place outside the realm of attention as defining theories of attention suggest (e.g. Broadbent, 2013; Treisman, 1964; Deutsch & Deutsch, 1963). Therefore, the majority of research presented here is based on experiments investigating passive perceptual processing.

All experimental paradigms designed to investigate prediction in auditory perception build up a form of auditory predictability and then observe whether the processing of upcoming stimuli is modulated by the predictive information. There are, however, different approaches and a variety of experimental paradigms designed to examine those questions (see also Bendixen et al., 2012).

### 1.3.1 Correlates of predictive auditory processing in response to rule-violation

A common way of investigating auditory perceptual processing and predictive mechanisms within auditory perception is by presenting acoustic stimuli that follow certain regularities based on acoustic features (e.g. frequency, intensity, location, timbre or presentation rate). In the experimental condition, a violation of these rules is invoked by changing a critical feature which dictates the regularity. An example of such an experimental design, often referred to as oddball paradigm, is depicted in Figure 1.2 (right panel). Researchers use this technique to contrast ERPs in response to standard stimuli (i.e. stimuli in compliance with the rule) and ERPs in response to deviant stimuli (i.e. stimuli violating the rule). The differences observed between these electrophysiological brain responses might provide helpful information about the underlying processes, and when measured under properly controlled conditions, they can be a useful tool to investigate the predictive nature of auditory perception.

One such ERP component is the MMN which has been used as a tool in a wide range of different research questions (for reviews see: Winkler, 2007; Kujala, Tervaniemi, & Schröger, 2007). The MMN is typically elicited by deviant tones (as described above), peaks at around 100-250 ms relative to stimulus-onset, has a frontocentral negative scalp distribution and is assumed to be generated in auditory areas of supratemporal cortex (Giard, Perrin, Pernier, & Bouchet, 1990; Molholm, Martinez, Ritter, Javitt, & Foxe, 2005; Sabri, Liebenthal, Waldron, Medler, & Binder, 2006). The MMN can be observed best when plotted as the difference wave of the deviant ERPs and the standard ERPs. An exemplary illustration of the MMN component is depicted in Figure 1.2 (left). The MMN is often interpreted in the context of prediction because it is assumed to be a correlate of the prediction error that arises from the comparison between top-down prediction and bottom-up sensory activation (Schröger, 2007). Alternatively, it has been suggested that the MMN reflects the update of a perceptual model of the external environment (Winkler & Czigler, 1998; Winkler, Karmos, & Näätänen, 1996). Opposing alternative explanations claim that the MMN is generated by so-called fresh-afferent neural activity towards deviating information (May & Tiitinen, 2010). However, MMN can not only be elicited by rule violations, like unexpected changes in the physical properties of a tone but also by violations of abstract rules (driven by relative relations between certain stimuli). For instance, MMN can be measured in response to the violation of a relational rule that links tone duration with tone frequency within a sequence of consecutively presented tones. For example, long tones that are followed by high tones and short tones that are followed by low tones (Bendixen et al., 2008; Paavilainen, Arajärvi, & Takegata, 2007). These results suggest that the generation of MMN is not merely based on different states of refractoriness. Furthermore, it has been argued that the MMN does not necessarily reflect predictive processes but might work in a retrospective manner (Schwartze, Tavano, Schröger, & Kotz, 2012) which implies that the MMN is the outcome of a matching processes that takes place after stimu-

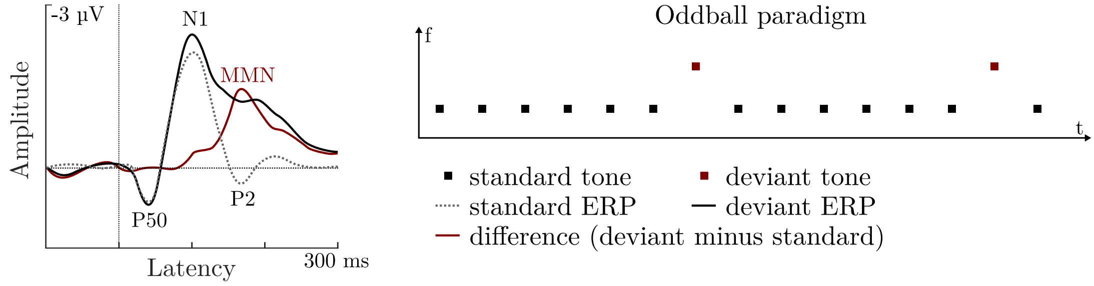


Figure 1.2: Right: a common implementation of the oddball paradigm. Tones of a certain frequency are presented isochronously (standard tones) and occasionally a tone with different frequency occurs (deviant tones). Left: exemplified ERP responses. Both, ERPs in response to standard and deviant tones usually display an early positive deflection at 50 ms (P50) followed by a negative peak at around 100 ms (N1). ERPs in response to deviants usually display an enhanced N1 amplitude. Additionally an MMN is elicited by deviant tones which can be seen as a negative deflection following the N1 in the difference wave between deviant and standard tones.

lus encoding without any contributions of prediction. However, other studies using mismatch paradigms showed that neural mechanisms detecting mismatch might indeed be based on predictive mechanisms. Grimm, Escera, Slabu, and Costa-Faidella (2011) showed with an oddball paradigm with standards and occasional frequency deviants that not only the MMN but also the Nb-component of the middle latency auditory-evoked responses was enhanced by deviant tones. Those responses are considered the earliest auditory cortex responses, usually observed between 20-80 ms (Yvert, Crouzeix, Bertrand, Seither-Preisler, & Pantev, 2001). Such early effects suggest that predictive and not just retrospective mechanisms are involved in mismatch detection.

Other ERP components indicate that auditory predictions might also be driven by visual information. A so-called incongruency response (IR) is elicited when a prediction about an upcoming tone is presented visually (e.g. notes on a screen) but the actual tone differs in pitch from what was visually predicted (Widmann, Gruber, Kujala, Tervaniemi, & Schröger, 2007). The authors report a negativity in the ERP observed around 100–140 ms relative to stimulus-onset, which probably indicates the detection of mismatch between the visually-based auditory prediction and the actual auditory stimulus.

There are furthermore several ERP components indicating predictive processing of more complex acoustic signals like speech and music. For example, the early right-anterior negativity (ERAN) is observed in response to musical violations

(Koelsch, Gunter, Friederici, & Schröger, 2000). In the domain of language, the early left-anterior negativity (ELAN) has been observed in response to syntactic violations in the form of a left-anterior negative deflection in the ERP at around 150-200 ms after syntax violation (Hahne & Friederici, 1999). Later correlates of violation-detection are the N400, a negative deflection at approximately 400 ms relative to the onset of semantically unexpected stimuli (Kutas & Hillyard, 1980) and the P600 (Osterhout & Holcomb, 1992) which can typically be observed after 600 ms relative to a phrase structure violations in spoken sentences (Friederici, 2002). All these indicators provide valuable insights into possible implementations of predictive auditory processing. However, it should be noted that all the electrophysiological correlates mentioned throughout this section are always measured after the onset of a critical event. Hence, it cannot be ruled out that these ERP components are evoked by processes that entirely work in a retrospective fashion without any contributions from predictive processing.

### 1.3.2 Correlates of predictive auditory processing in response to rule-confirmation

An alternative way of investigating predictive processes in audition is provided by so-called match paradigms. For instance Haenschel, Vernon, Dwivedi, Gruzelier, and Baldeweg (2005) sequentially presented identical acoustic stimuli for a certain amount of time until one feature of the tone (e.g. its frequency) changed, and again is repeated for a while. The authors found a positive deflection in the ERP on frontocentral electrodes ranging from 50 to 250 ms after tone onset which increased with the number of repetitions of the tones with identical frequency. This so-called repetition positivity (RP) is interpreted as a measure of stimulus adaptation due to the forming of sensory memory representations which might lead to an increasingly suppressed processing of the tones. It has furthermore been shown that this brain response is not only affected by the information of what is likely to happen next but also by information about when it is likely to happen (Costa-Faidella, Baldeweg, Grimm, & Escera, 2011). RP has also been investigated by Bendixen et al. (2008) using more complex and abstract rules (relational rules between certain stimulus features as described in Section 1.3.2). However, these results do not provide clear evidence regarding the predictive nature of auditory perception because they do not rule out the possibility that the match process indicated by the RP is carried out strictly retrospective without the use of any predictive mechanism.

Another investigation of match processes was carried out by means of time-frequency analyses of the oscillatory EEG activity (Schadow, Lenz, Dettler, Fründ, & Herrmann, 2009). It has been suggested that oscillatory activity in the gamma range (20-80 Hz) might be related to bottom-up and top-down processes in auditory perception (Herrmann, Munk, & Engel, 2004). This observation is called evoked gamma-band response (GBR) and it emerges quite early at around 40-100 ms after stimulus-onset (for a review, see Kaiser & Lutzenberger, 2005). Schadow et al. (2009) presented sequences of ascending or descending sine tones that were either presented regularly (six tones regularly ascending or descending) or an irregularity was introduced by a frequency deviant at either the third or fifth position of the sequence which represents a rule violation. The authors investigated the early evoked GBR by means of time-frequency analysis and found an enhanced evoked GBR as early as 50 ms after tone onset of regularity-conforming tones as compared to irregular tones. As for the RP, this finding still does not provide clear evidence in favor of predictive processing over a retrospective processing mode within perception but the early onset of the response argues in favor of the predictive account.

### 1.3.3 Correlates of predictive auditory processing in response to stimulus omissions

Another refined way of investigating predictive processes in audition, is to examine brain responses to acoustic events which were highly expected but then turn out to be absent. This provides means to examine the true electrophysiological correlates of prediction because the response is not superimposed with the sensory processing of an acoustic event (Bendixen et al., 2012). Using such a paradigm, Janata (2001) observed the N1 component of the ERP to be elicited without any physical input. The N1 peaks at around 80-120 ms and is influenced both by sensory (external) influences as well as internal influences like attention (Näätänen & Picton, 1987). Janata (2001) presented sequences of tones which were learned by the subjects. Later they measured the subjects' EEG while presenting incomplete tone sequences which should be ignored in one condition, while in another condition the subjects were asked to imaginarily continue the sequence. They found N1 components to imagined tones to equal components in response to actual tones. This was, however, only the case if the subjects imagined the tones but not if they were ignored.

Even though such paradigms provide better control regarding the influence of physical stimuli on the measured outcomes, problems still exist in the interpretation of experimental results with regard to the distinction between retrospective vs. predictive aspects of auditory perception. Bendixen et al. (2009) tried to overcome these limitations by means of a passive omission paradigm. Since the experimental logic of the current thesis is mainly based on this study, it will be introduced in greater detail. As can be seen in Figure 1.3, this was done by presenting a sequence of tones in which every other tone was identical to the preceding one (pairs of tones with identical frequency) and therefore was predictable, while the other half of the tones consisted in random frequency jumps and therefore were unpredictable. Occasionally, tones were omitted either at the first position (restorable condition) or the second position (predictable condition) of the tone pairs. In a control condition there were no frequency pairs at all and therefore no predictive information about the tones' frequencies was given. The authors observed that ERPs in response to omissions of predictable tones differed from omissions in the restorable and control conditions and resembled ERPs of tones up to 50 ms. This initial similarity in the processing of a predictable tone and of the omission of a predictable tone was taken to conclude that neural circuits are pre-activated if the context provides enough information about the forthcoming event.

There are many other studies using different approaches to examine predictive processes in auditory perception, e.g. investigations about the processing of self-generated sounds indicated by the N1 suppression effect (Bäss, Jacobsen, & Schröger, 2008; Martikainen, Kaneko, & Hari, 2005). However, these paradigms commonly require subjects to perform active tasks which introduces additional complex phenomena, like perception-action coupling. Since the scope of the current thesis covers perceptual processes per se, these studies will not further be discussed here. From the evidence accumulated so far, Bendixen et al. (2012) conclude that an auditory prediction mechanism might consist of three essential steps (see also Schröger, 2007): 1. relations of successive stimuli are extracted and represented, 2. those relations are compared and 3. based on this comparison, regularities can be extracted which might be used to form predictions about upcoming sensory input. Mathematical and computational models based on this design have been developed which further substantiate the plausibility of such a mechanism (Kiebel, von Kriegstein, Daunizeau, & Friston, 2009). Furthermore, using biologically realistic simulations of neural networks based on predictive mechanisms, it is possible to replicate a lot of findings as discussed in Section 1.2.

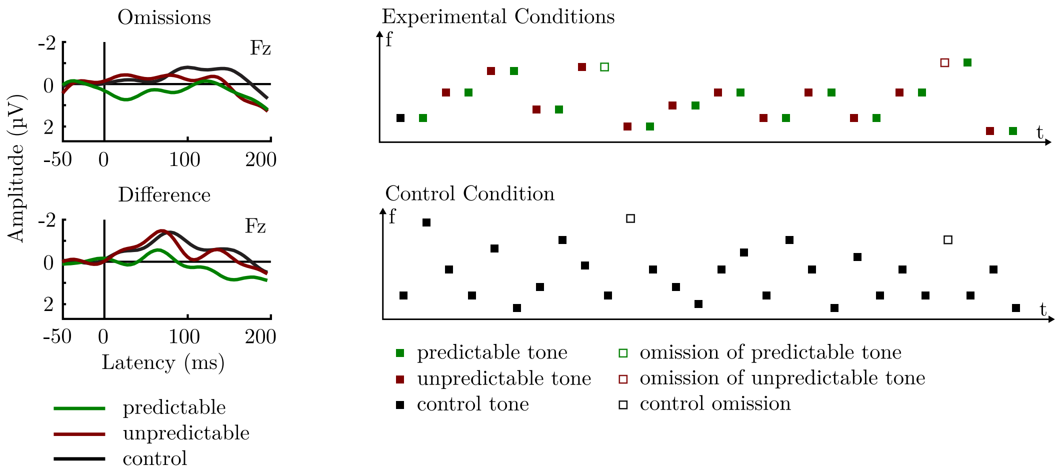


Figure 1.3: Experimental paradigm and electrophysiological results of a study by Bendixen et al. (2009) on which the current thesis is based. They used an omission paradigm to investigate whether brain responses to omissions of predictable tones differ from responses to omissions of unpredictable tones. Unlike illustrated here, the omissions of predictable vs. restorable tones were not presented within the same sequence but in different experimental blocks. The authors observed that ERPs in response to omissions of predictable tones differed from omissions in response to unpredictable tones in both the experimental and the control condition. The difference waves resemble the difference between ERPs in response to tones and ERPs in response to omissions in the respective conditions. Omission ERPs of predictable tones resembled ERPs of tones up to 50 ms after stimulus-onset.

## 1.4 Limitations of previous research and aims of the thesis

Bendixen et al. (2009) argue that a lot of studies discussed so far may be in favor of the predictive nature of auditory perception but do not rule out the retrospective account. That is, the system may as well aim at matching each incoming stimulus to the preceding stimuli only after it has encountered the stimulus and not actively make predictions about forthcoming stimuli based on a model derived by the preceding sequence. They therefore attempted to distinguish between the retrospective and prospective (predictive) account of perception by the technique of omitting highly predictable tones to investigate brain responses of these events without the interference caused by the processing of the tones themselves (see also Section 1.3.3). Since this paradigm introduces some major improvements with respect to the conclusions that can be drawn from the obtained electrophysiological results, the basic logic of Bendixen et al. (2009) was adopted in the current thesis.

However, some problems still remain. In the original study the reliability of the predictive information is always 100 %. There are only tone pairs in the experimental conditions, so the system can be sure about the frequency of the second tone before encountering it. In the real world however, this reliability is rarely given. Regularities in the natural world almost always emerge with a certain amount of uncertainty. This is due to the mere fact that almost no object in the real world "behaves" (i.e. emits auditory signals) in a 100 % regular manner, and even if it does, sensory information from other sources can interfere with this regularity. If this paradigm indeed taps into a mechanism that is generally embedded in our perception system, then the observed effects should adapt to a certain amount of uncertainty of predictive information. In other words, the predictive processing effect should still be observed if the prediction holds in only, e.g. 75 % of the cases.

Another issue concerning the investigation of predictive processing with electrophysiological measures is that no hypotheses exist about the polarity of prediction-related ERP effects. As described above, previous studies usually contrasted ERPs in response to fully predictable events with ERPs in response to unpredictable events. Any differences in the ERPs in response to these two extreme conditions (whether positive or negative in polarity) are then taken as electrophysiological correlates of predictive processing in the brain. In order to decrease the likelihood of incidental findings, it is important to probe more than two conditions, creating the opportunity to find links between a gradual variation of predictive relations and correspondingly graded ERP effects. Therefore, the current thesis aimed at implementing such a gradual manipulation of predictability which enables to systematically investigate auditory predictive processing under uncertainty. The experimental manipulations comprised extreme conditions (fully predictable versus fully unpredictable stimulus arrangements) as well as several intermediate levels of predictive relations. Predictability was created by means of tone frequency repetition, based on the paradigm by Bendixen et al. (2009). Departing from previous designs, in one Experiment frequency repetition (and hence predictability) occurred with different degrees of reliability to further investigate to what extent the auditory system adapts to different degrees of uncertainty in the sensory input. In another Experiment different degrees of accuracy of predictive relations between tones were introduced to further characterize the fault tolerance of the proposed predictive auditory mechanism. The rationale of applying two different ways of introducing uncertainty into the formation of predictive relations aimed at cross-validating the findings as a measure against incidental observations.

Another major pitfall of previous investigations is that processing differences between predictable and unpredictable events have always been shown after stimulus-onset. However, as soon as the stimulus has been presented, not only prediction but also prediction error signals come into play, and it is highly difficult to separate the two (SanMiguel, Widmann, Bendixen, Trujillo-Barreto, & Schröger, 2013; Yordanova, Kolev, & Kirov, 2012). An elegant way to demonstrate that prediction processes are at work would be to demonstrate ERP modulations shortly before the onset of predictable (as opposed to unpredictable) events.

If the auditory system processes its input in a predictive manner, systematic ERP modulations by the reliability and accuracy manipulations should be observed in the current thesis. If these ERP effects appeared before rather than after the onset of a predictable tone, this would provide compelling evidence that the brain engages in predictive processing in a literal sense. Finding graded effects of the predictability manipulations would further demonstrate that sensory predictions flexibly adapt to the certainty with which predictions can be made. The relevant time-ranges and polarities of these graded ERP effects would provide valuable insights for computational models of predictive processing (Deneve, 2008; Kiebel et al., 2009; Rao & Ballard, 1999; Srinivasan et al., 1982; Wacongne et al., 2012) especially if they occurred consistently across both predictability manipulations.

The following part of the current thesis focuses on the influence of a gradual manipulation of uncertainty on "traditional" (post-stimulus) ERP results following Bendixen et al. (2012). After that, the results obtained so far were investigated with the aim of identifying prediction-related pre-stimulus ERPs. Two more Experiments were introduced to further investigate the characteristics of pre-stimulus ERP effects associated with predictive processing. In the final part, both, post-stimulus and pre-stimulus effects are re-investigated after combining the results of several Experiments to overcome certain methodological pitfalls and to gain a more precise view on the pattern of results.

# 2 Traditional correlates of auditory prediction

As introduced in Section 1.3, previous investigations of predictive auditory processing usually build up some form of regularity based on which the system might predict future incoming events. Different paradigms have been developed to investigate proposed ERP-correlates of prediction and prediction error signals, like e.g. the MMN, N1, ELAN or the N400 using mismatch paradigms, correlates like the RP as well as N1 attenuation within match-paradigms or by investigating a variety of such electrophysiological correlates within omission paradigms. All these measures usually occur after stimulus-onset and therefore are considered traditional correlates of auditory predictive processing. In this Section of the thesis, two experiments will be introduced which use different kinds of gradual manipulations of predictive relations between stimuli in order to investigate such traditional correlates of predictive auditory processing.

As mentioned before, the basic logic of the experiments was adopted from Bendixen et al. (2009) which rapidly presented isochronous tone sequences as described in Section 1.3.3. However, using such rapid presentation rates introduces a constraint as to which ERP components can be investigated. In general, rapid stimulation rates render it very difficult to differentiate whether the measured responses are indeed elicited by the current stimulus or whether they are driven by later processing stages of previous stimuli. Furthermore, some components of the ERP are strongly affected by the rate at which stimuli are presented. For example, the N1 is known to typically show a strong decrease in amplitude when stimuli are rapidly presented. According to Budd, Barry, Gordon, Rennie, and Michie (1998), this process has traditionally been ascribed to either habituation processes (Sokolov, 1963) or to refractory processes linked to the recovery cycle of underlying neural generators (Ritter, Vaughan, & Costa, 1968; Callaway, 1973). Therefore, the following investigation focuses only on very early correlates of auditory prediction which are less prone to such strong influences of the presentation rate, like the P50 component or early parts of the RP.

## 2.1 Experiment 1<br>Reliability of auditory predictions

The first experiment was conducted at the Cognitive and Biological Psychology Lab of the Department of Psychology (University of Leipzig). It was aimed at systematically investigating the influence of repetition reliability on the processing of omissions of potentially predictable vs. unpredictable tones by means of ERPs. Data acquisition and parts of the analysis were carried out within the context of my master thesis, however within the current thesis the data was analyzed more comprehensively which yielded new and more extensive results. The experiment generally followed the logic of Bendixen et al. (2009). Additionally, the reliability of auditory predictive relations was varied by manipulating the conditional probability of frequency repetitions within isochronous tone sequences across five conditions. Repetition reliability ranged from 0 % (no repetitions at all) to 100 % (fully predictable repetitions in every other tone in the sequence) in steps of 25 % (resulting in the conditions: 0 %, 25 %, 50 %, 75 %, 100 %). As in the original study, an effect of reliability of predictive relations is expected between the extreme conditions (0 % vs. 100 % Repetition Reliability). If the observed mechanism really represents a general and adaptive property of auditory perception, the effect should systematically grow depending on the certainty of the predictive information. Either a linear trend in the size of the effect should be observed, or the effect should start to show up after a threshold (minimum probability for which it "pays off" to make predictions) has been surpassed. In the latter case, the system could be said to act according to an all-or-none principle.

Besides testing the effect of conditional probability on the observed ERP modulations, the design of the present study may introduce another advantage over the original study (Bendixen et al., 2009). In the original study, the comparison was based on experimental and control conditions that were physically very different from each other (cf. Figure 1.3). Predictable tones in the experimental condition are repetitions of the previous ones, whereas tones in the control condition are always of a different frequency than the ones before. Because physical stimulus repetition vs. change has been shown to dramatically alter neural responses (see e.g. Ritter et al., 1968; Sams, Alho, & Näätänen, 1984), the tones in the experimental and the control conditions of the original study might in part not be comparable. While it is valid to compare tones in the first position of a pair (restorable) with tones in the control condition (both of them constituting physical changes), it can be problematic to compare tones in the second position in a pair (predictable) with the control trials because one of the events is a repetition, while the other is a change. It might therefore be that these tones are not only processed differently because of different degrees of predictability but also because of different states of refractoriness of the involved neuronal populations. The current thesis inherently provides means to cope with this problem because in the conditions with 25 %, 50 % and 75 % Repetition Reliability, some of the tones are (to some degree) predictable and indeed constitute physical repetitions, while other tones are (to the same degree) predictable but then constitute physical changes. A systematic comparison of these event types will allow for a well-controlled comparison of the tone ERPs in addition to the omission-related ERPs. Thus, in contrast to the original study, the current investigation not only analyzes potential effects of predictability within omitted tones but also within the (standard) tones themselves. As was mentioned before, measures of perceptual prediction should not only show up in situations where predictions are violated but also in situations in which predicted events actually occur like in the case of the RP (Haenschel et al., 2005). The present investigation thus enlarges the previous approach in several theoretically and methodologically important ways.

### 2.1.1 Methods

**Subjects** Twenty healthy subjects (16 female, 19-43 years old, mean age: 24.7 years) participated in the experiment. Prior to the experiment, subjects were asked to give written consent in accordance with the Declaration of Helsinki (World Medical Association, 2013) after being informed about the nature of the experiment. Subjects received course credit or modest financial compensation for their participation. Subjects with less than 80 % artifact-free data were rejected from further analysis in all of the experiments. In Experiment 1 the average proportion of artifact-free data was 95.41 % (standard deviation [SD]: 4.33 %, minimum: 84.56 %). Therefore, data of all subjects was used for further analysis. Nineteen of the twenty subjects were right handed (mean laterality index: 86.5) and one was ambidextrous (laterality index: -14.3) according to a German version of the Edinburgh Handedness Inventory (Oldfield, 1971).

**Experimental procedures** Subjects were seated in a comfortable chair inside an acoustically attenuated and electrically shielded testing chamber (IAC Acoustics, Niederkrüchten, Germany). Isochronous tone sequences were presented binaurally via headphones (Sennheiser HD25-1, 70Ω) with a level of 70 decibel (dB) sound pressure level (SPL). Participants were instructed not to pay attention to the tones while watching a self-selected, silenced movie with subtitles on a screen positioned outside of the testing chamber, visible through a glass pane. Tones were synthesized with Matlab R2011b (The MathWorks Inc., Natick, USA) and presented using the Psychophysics Toolbox extension for Matlab (Brainard, 1997). Tones were of 50 milliseconds (ms) duration (half-raised cosine onset and offset ramps of 5 ms each) and presented with a stimulus-onset asynchrony (SOA) of 150 ms. Each tone was assigned a frequency between 400 - 1000 Hz. Following the condition-specific constraints described below, this frequency was either predictable on the basis of the preceding tone, or it was randomly chosen with the restriction that the frequency of the preceding tone was at least one semitone (5.9 %) apart.

**Experimental paradigm** In five conditions, the conditional probability for a tone to be a repetition of the previous one (i.e. forming a tone pair) was 0 %, 25 %, 50 %, 75 % and 100 % as displayed in Figure 2.1. After each repetition, a frequency change was enforced (i.e., there were never two frequency repetitions in a row). As a consequence, every tone in condition 1 (0 %) was randomly chosen (hence unpredictable) with respect to its frequency, whereas every other tone in condition 5 (100 %) was fully predictable. The intermediate conditions 2 to 4 were partly predictable with different degrees of repetition reliability. Tones that were unpredictable with respect to their frequency (the first tones of the frequency pairs) will be called *certainly unpredictable* and tones that could potentially be predicted based on the frequency of the previous tone (second tones of the pairs) will be called *potentially predictable* throughout the whole thesis. In 5 % of the cases, tones were replaced with a 50 ms gap (silence). Those omissions were presented at random positions with the restriction that two omissions were at least 1050 ms apart. 2.5 % of the omissions were presented at positions of certainly unpredictable tones and the other 2.5 % of omissions were presented at positions of potentially predictable tones in conditions 2 to 5. In condition 1, all the omissions were presented at positions of certainly unpredictable tones. In condition 1, 4000 stimuli were presented of which 200 were omissions; in conditions 2 to 5, 8000

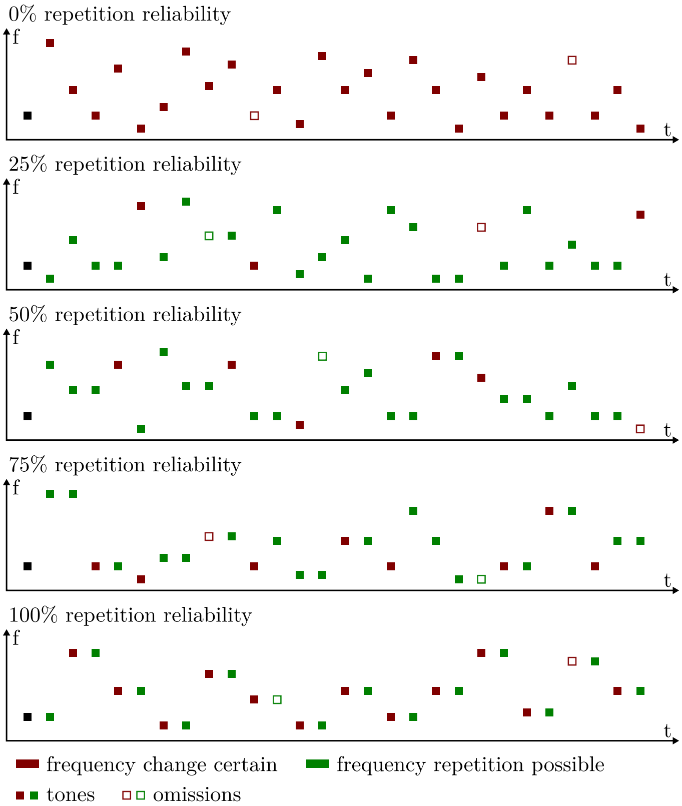


Figure 2.1: Design of Experiment 1: manipulating the repetition reliability of predictive relations between successive tones. The conditional probability of frequency repetition was varied in five conditions (0 %, 25 %, 50 %, 75 %, 100 %). Positions where a frequency change was certain are marked in red; positions where a frequency repetition was possible are marked in green. 5 % of tones were replaced by omissions (equally distributed across certainly unpredictable and potentially predictable tones). Omissions were presented in random order.

stimuli were presented with 200 omissions at positions after a frequency repetition and another 200 omissions at positions after a frequency change. The whole experiment consisted of 18 blocks with 2000 trials each. Therefore, there were two blocks for condition 1 and four blocks for conditions 2 to 5 each. All blocks of one condition were presented consecutively. Condition order was counterbalanced between subjects. Net experimenting time was approximately 90 minutes. Together with electrode application and removal as well as breaks between the experimental blocks, the overall duration of the experiment amounted to 3.5 hours.

**Electrophysiological data acquisition** EEG was measured using a BioSemi ActiveTwo System (BioSemi, Amsterdam, The Netherlands) with active Ag/AgCl electrodes from 64 scalp positions according to the 10-10 extension of the International 10-20 System (American Electroencephalographic Society, 1994) and two further electrodes at the left and right mastoid (M1, M2). The horizontal Electrooculogram (EOG) was measured with electrodes placed at the outer canthi of the left and right eye. The vertical EOG was obtained from separate electrodes placed above and below the left eye. The reference electrode was placed at the tip of the nose. EEG and EOG signals were amplified and recorded with a sampling rate of 512 Hz. An online lowpass filter of 250 Hz was applied to the raw data to avoid aliasing (Nyquist, 1928).

**Electrophysiological data analysis** EEG data were analyzed offline using the EEGLAB toolbox for Matlab (Delorme & Makeig, 2004). To enhance conformity between all experiments of the current thesis, the sampling rate was reduced to 500 Hz. An independent component analysis (ICA) approach was used to remove artifacts and to increase signal-to-noise ratio (SNR). ICA artifact correction was implemented by first concatenating the raw data of each subject and high-pass filtering (-6 dB cutoff: 1 Hz, transition width: 1 Hz, order: 1812) the data with a zero-phase Kaiser windowed sinc finite impulse response (FIR) filter (maximum passband deviation: 0.1 %, stopband attenuation: -60 dB, Kaiser Beta: 5.65326). The filtered data was segmented into epochs of 1 second length. Epochs containing unique, non-stereotyped artifacts were rejected. Independent Component Analysis decomposition was applied on the remaining data using the extended infomax ICA algorithm of Bell and Sejnowski (1995). In order to identify artifact-related component activity, in a second step the raw data was again high-pass (-6 dB cutoff: 0.1 Hz, transition width: 1 Hz, order: 1812) and low-pass filtered (-6 dB cutoff: 48 Hz, transition width: 1 Hz, order: 1812) using a zero-phase Kaiser windowed sinc FIR filter (as described above). ICA activity related to eye movements, eye blinks, cardiac signals, muscle noise, and line noise were removed (e.g. Jung et al., 2000). Channels with technical malfunction were interpolated using spherical spline interpolation (Perrin, Pernier, Bertrand, & Echallier, 1989). The data were epoched from -150 to 150 ms relative to stimulus-onset. Epochs with EEG or EOG changes exceeding 100 µV rejected from further analysis leading to an average of 4.59 % data loss. In line with Bendixen et al. (2009), ERPs were baseline corrected using the pre-stimulus time window as baseline window (-150 ms to 0 ms relative to stimulus-onset). For averaging of the tone ERPs all omissions were excluded. Furthermore, tones following an omission within 600 ms of omission-onset were also excluded. That led to the exclusion of 800 tones in condition 1 and 1600 tones in Condition 2 to 5. Grand-average ERPs were computed for all tone and omission types separately for each condition. Statistical analyses were consistently carried out on ERPs obtained from electrode position Cz (central midline electrode placed above the vertex). This electrode position was chosen for two reasons. First, because ERPs with contributions from primary auditory cortex usually show a frontocentral scalp distribution (see e.g. Haenschel et al., 2005; Winkler, 2007; Näätänen, Paavilainen, Rinne, & Alho, 2007) and second, because different electrode layouts were used across all experiments and electrode position Cz was available in all the layouts that were used.

First, in order to test whether the findings of Bendixen et al. (2009) can be replicated, a within-subject analysis of variance for repeated measures (RMANOVA) was conducted with the factor Stimulus Type (2 levels: change certain, repetition possible) and the factor Repetition Reliability (5 levels: 0 %, 25 %, 50 %, 85 %, 100 %) for omissions measured at Cz in the interval of 0 - 50 ms relative to stimulus-onset. Note, in condition 1 there were only omissions of certainly unpredictable tones, hence the same stimuli were used within the pool of certainly predictable and potentially unpredictable tones in condition 1. Note also that, in contrast to the original study by Bendixen et al. (2009), the time range for obtaining average amplitude values started at stimulus-onset and not 10 ms after stimulus-onset. This approach was applied because a "real" prediction response should be present with (or even slightly before) stimulus-onset (Wacongne et al., 2012).

To better distinguish between the different stimulus types in Experiment 1 and to better convey the rationale of the second analysis, the tones were assigned to three different categories. As can be seen in Figure 2.2, tones either occurred after a frequency repetition or after a frequency change. The tone frequency always changed after every repetition, hence these events are called "tones - change certain" (a). When tones occurred after a frequency change, they could constitute a frequency repetition of the previous tone. If this was the case, these tones would form the category b) "tones - repetition possible - repetition occurred". If, on the other hand, a tone after a frequency change was again itself a frequency change relative to the previous tone, this would fall into category c) "tones - repetition possible - change occurred". There were no repetitions at all in condition 1 and tone frequencies were always changing. Therefore, all tones in this condition are considered a) "tones - change certain". In condition 5, only tones of category a) and b) were possible because the whole sequence was composed of tone pairs. Thus, it is not

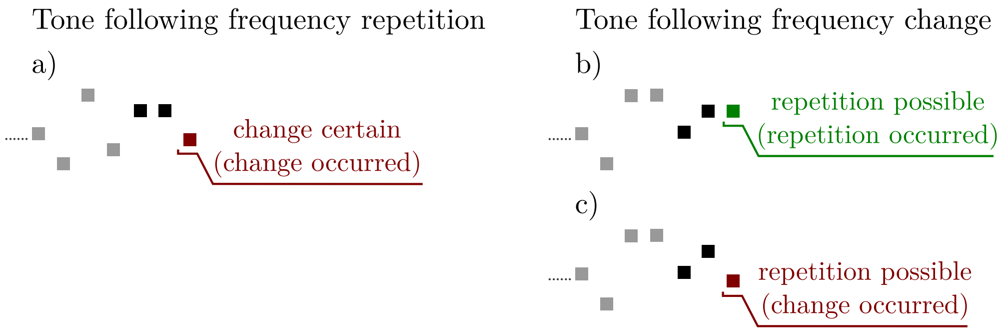


Figure 2.2: Different tone types: Tones following a frequency repetition certainly change in frequency (a). Tones following a frequency change can be a frequency repetition of the previous tone (b) or they can change in frequency (c). Therefore tones a) and c) are physically identical whereas tone b) differs from them.

possible that a tone changes its frequency after a frequency change occurred in the previous one already. In the remaining conditions 2 to 4, all three event types occurred (with different proportions depending on the conditional probability of tone repetition, see Table 2.1).

The second analysis aimed at distinguishing different influences (repetition reliability vs. stimulus type [b vs. c]) on the potentially predictable tones themselves. A further goal was to characterize early vs. late contributions to prediction-related responses. Therefore a within-subject RMANOVA was conducted with the factors Stimulus Type (2 levels: potentially predictable - change occurred vs. potentially predictable - repetition occurred), Window (2 levels: early vs. late) and Repetition Reliability (4 levels). By design, there were no tones of category b) in condition 1, and in condition 5 there were no tones of category c). Therefore, condition 1 to 4 (0 % to 75 %) was used for potentially predictable tones that constituted a frequency change (a and c) and condition 2 to 5 (25 % to 100 %) was used

Table 2.1: Numbers of stimuli per condition in Experiment 1. Displayed are the numbers of stimuli, organized in categories a) change certain - change occurred, b) repetition possible - repetition occurred and c) repetition possible - change occurred per level of Repetition Reliability.

| | condition | | | | |
|---|---|---|---|---|---|
| | 0 % | 25 % | 50 % | 75 % | 100 % |
| category | | | | | |
| a) | 3800 | 800 | 1800 | 2800 | 3800 |
| b) | 0 | 980 | 1940 | 2880 | 3800 |
| c) | 0 | 5820 | 3860 | 1920 | 0 |

for potentially predictable tones that constituted a frequency repetition (b). The early interval was set from 0 ms to 75 ms and the late interval from 75 ms to 150 ms relative to stimulus-onset. These intervals were objectively chosen to represent the first and second half of the post-stimulus interval and to be of equal length.

To study the topographical distribution of the ERP effects across conditions, scalp potential maps of the ERPs in the extreme conditions of each experiment were created in the respective analysis interval. For all RMANOVAs, the Greenhouse-Geisser correction was applied and the epsilon correction factor is reported whenever the sphericity assumption was violated as indicated by Mauchly's test.

### 2.1.2 Results

Grand-average ERPs of the omissions are shown in figure 2.3. Scalp topographies of the omission ERP modulation by repetition reliability are shown in Figure 2.4. The 2 x 5 RMANOVA with the factors Stimulus Type and Repetition Reliability for tone omissions yielded a significant interaction of Stimulus Type by Repetition Reliability [F(4,76)=7.247, p=0.00005, $\eta^2$=0.276, $\epsilon$=0.028]. Furthermore, there was a significant main effect of Stimulus Type [F(1,19)=36.353, p=0.00001, $\eta^2$=0.657] but not for Repetition Reliability [F(4,76)=1.043, p=0.39092, $\eta^2$=0.052, $\epsilon$=0.018].

As a follow-up analysis, one-way RMANOVAs with the factor Repetition Reliability were conducted for each level of the factor Stimulus Type. The RMANOVA for omissions of certainly unpredictable tones yielded no main effect of Repetition Reliability [F(4,76)=1.666, p=0.16660, $\eta^2$=0.081]. For omissions of potentially predictable tones on the other hand, there was a significant main effect of Repetition Reliability [F(4,76)=5.553, p=0.00056, $\eta^2$=0.226 (Bonferroni-corrected)]. This main effect furthermore followed a linear trend [F(4,19)=15.062, p=0.00101, $\eta^2$=0.442] which indicates more positive amplitudes associated with higher repetition reliability.

Grand-average ERPs (change occurred vs. repetition occurred) are shown in figure 2.5. Scalp topographies of the respective Repetition Reliability contrasts for potentially predictable tones (change occurred vs. repetition occurred) are displayed in Figure 2.6. The 2 x 2 x 4 within-subject RAMNOVA with the factors Stimulus Type, Window and Repetition Reliability for ERPs of potentially predictable tones revealed significant interactions of Stimulus Type by Window [F(1,19)=22.364,

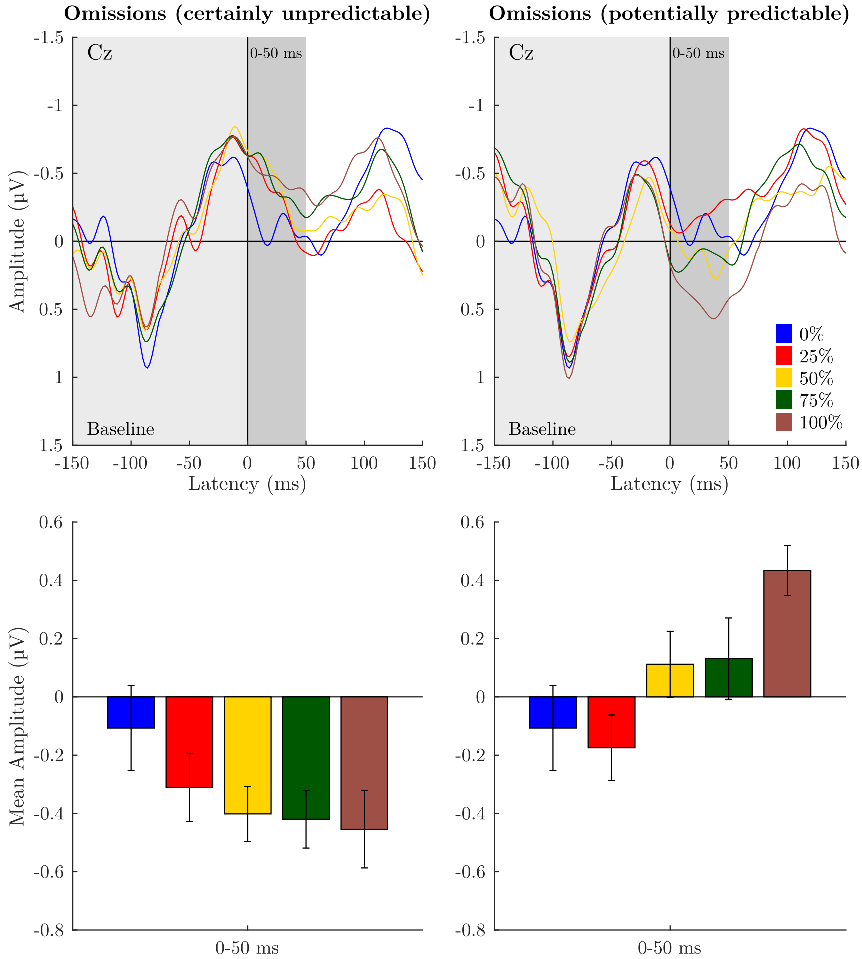


Figure 2.3: Electrophysiological Results. Upper panel: grand-average ERPs of omissions (certainly unpredictable vs. potentially predictable) across all levels of Repetition Reliability. Lower panel: bar diagrams with mean amplitude values for each condition in the interval of 0 ms to 50 ms relative to stimulus-onset. Error bars indicate standard errors of the mean.

p=0.00015, $\eta^2$=0.541], Stimulus Type by Repetition Reliability [F(3,57)=7.575, p=0.00024, $\eta^2$=0.285] and Window by Repetition Reliability [F(3,57)=4.829, p=0.00459, $\eta^2$=0.203]. Furthermore, there were significant main effects of Stimulus Type [F(1,19)=27.716, p=0.00004, $\eta^2$=0.593], Window [F(1,19)=33.791, p=0.00001, $\eta^2$=0.640] and Repetition Reliability [F(3,57)=14.641, $p<0.00001$, $\eta^2$=0.435]. However, a three-way interaction of Stimulus Type by Window by Repetition Reliability was not present [F(3,57)=1.475, p=0.23095, $\eta^2$=0.072].

Two-tailed paired t-tests were calculated as a follow-up analysis to resolve the Stimulus Type by Window interaction. Mean amplitudes did not differ statistically in the early window when averaged across all levels of Repetition Reliability

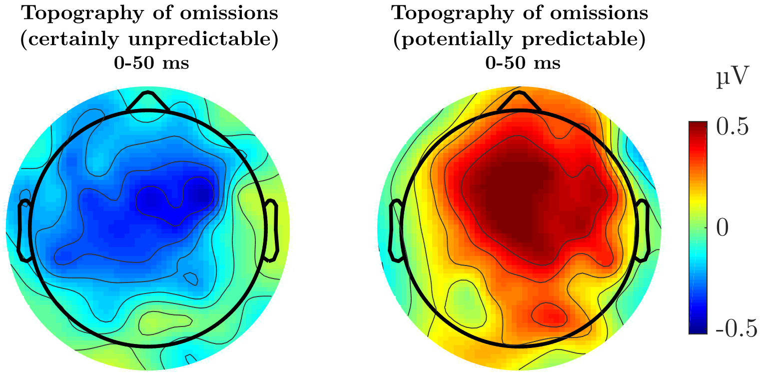


Figure 2.4: Topographical scalp voltage distribution of omission ERP modulation by repetition reliability in the interval of 0 ms to 50 ms relative to stimulus-onset. Plotted are the difference waves of the extreme conditions (100 % minus 0 %).

[t(19)=2.3, p=0.03298 (Bonferroni-corrected)]. In the late window however, there was a significant difference between events where a frequency change occurred in contrast to events where a frequency repetition occurred [t(19)=5.637, p=0.00002 (Bonferroni-corrected)]. This difference indicates generally more positive amplitudes when a change occurred vs. when a repetition occurred irrespective of the Repetition Reliability. A one-way within-subject RMANOVA with the factor Repetition Reliability was conducted for each level of the factor Stimulus Type averaged across both windows. There was no main effect of Repetition Reliability for tones that changed in frequency across both windows [F(3,57)=4,169, p=0,03170, $\eta^2$=0.18, $\epsilon$=0.546 (Bonferroni-corrected)]. However, for tones that constitute a frequency repetition, there was a significant main effect of Repetition Reliability [F(3,57)=15.047, p<0.00001, $\eta^2$=0.442 (Bonferroni-corrected)] which followed a linear trend [F(1,19)=34,654, p=0.00001, $\eta^2$=0.646], indicating more positive amplitudes associated with higher degrees of Repetition Reliability for frequency repetitions across both windows. Finally, a within-subject RMANOVA with the factor Repetition Reliability was conducted for each window, averaged across Stimulus Type, to clarify the Repetition Reliability by Window interaction. For each window, there was a main effect of Repetition Reliability [early: F(3,57)=21,523, p<0.00001, $\eta^2$=0.531; late: F(3,57)=3,861, p=0.01388, $\eta^2$=0.169; both Bonferroni-corrected] which both followed a linear trend [early: F(1,19)=49,172, p<0.00001, $\eta^2$=0.721; late: F(1,19)=7,734, p=0.01191, $\eta^2$=0.0.289]. These results indicate a general effect of Repetition Reliability, both in the early and late window even though the effect is much smaller in the late time window ($\eta^2_{early} = 0.531$ vs. $\eta^2_{late} = 0.169$).

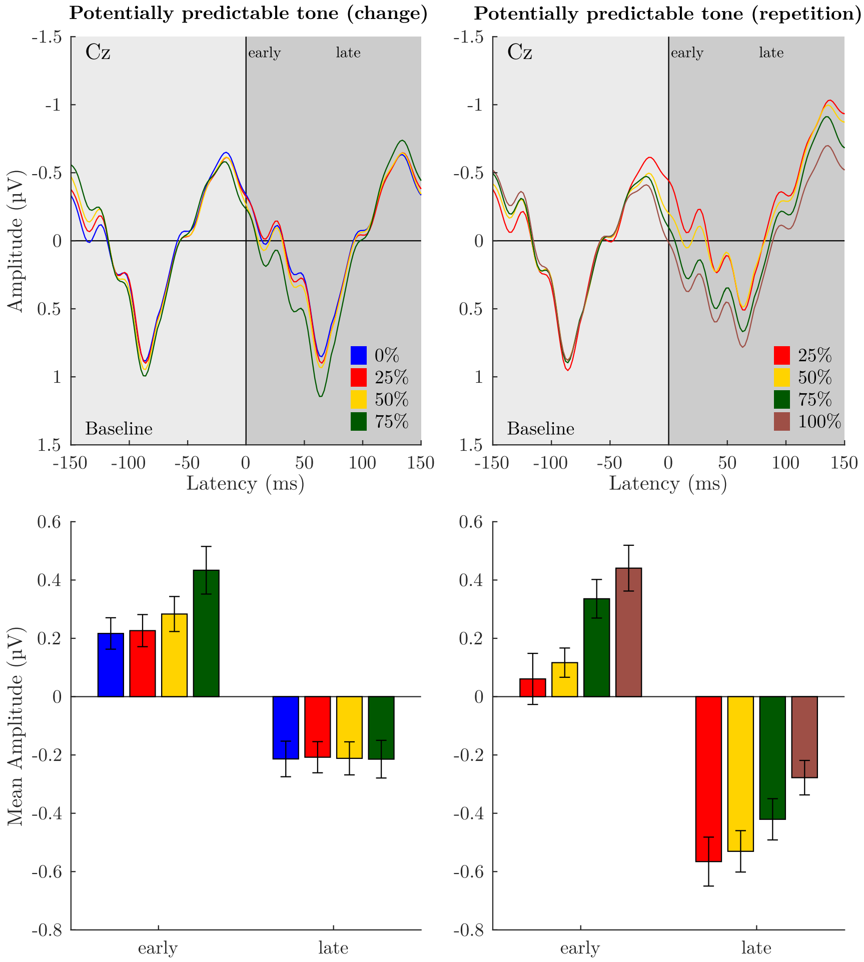


Figure 2.5: Electrophysiological Results. Upper panel: grand-average ERPs of potentially predictable tones (change occurred vs. repetition occurred) across selected levels of Repetition Reliability. Lower panel: bar diagrams with mean amplitude values for each condition in the interval of 0 ms to 75 ms (early window) and 75 ms to 150 ms (late window) relative to stimulus-onset. Error bars indicate standard errors of the mean.

### 2.1.3 Discussion

Predictive relations between subsequently presented tones varied systematically by manipulating the repetition reliability in five conditions from unpredictable (0 %) to predictable (100 %). The experiment was designed to investigate whether the proposed predictive auditory mechanism flexibly adapts to different degrees of predictive certainty embedded in the sensory environment. Evidence in favor of such flexible adaption was expected as a graded effect of repetition reliability in the ERPs in response to omissions of potentially predictable tones (but not of certainly unpredictable tones) immediately following stimulus-onset.

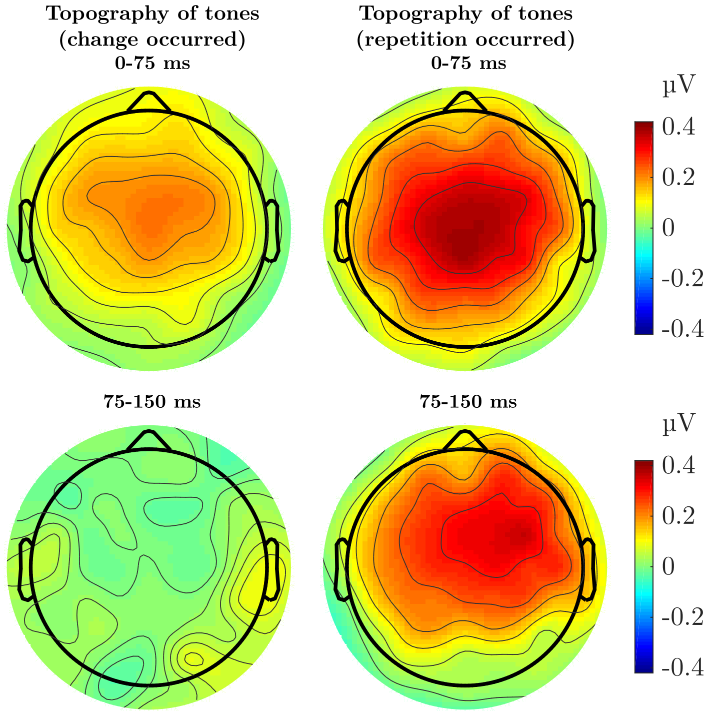


Figure 2.6: Topographical scalp voltage distribution of tone ERP modulation by repetition reliability of tones constituting a change (left) and tones constituting a repetition (right) in the interval of 0 ms to 75 ms (early window; top) and 75 ms to 150 ms (late window; bottom) relative to stimulus-onset. Plotted are the difference waves of the extreme conditions (75 % minus 0 % for frequency changes and 100 % minus 25 % for frequency repetitions).

ERPs in response to omissions of potentially predictable tones systematically diverged during the first 50 ms following stimulus-onset as evidenced by a significant interaction of stimulus type by Repetition Reliability in the omnibus-ANOVA. This was furthermore confirmed by a significant main effect of Repetition Reliability in a follow-up RMANOVA for potentially predictable tones. Furthermore, this effect followed a linear trend, indicating an increasingly positive deflection of ERPs in the initial 50 ms following stimulus-onset. However, no such effect was observed for omissions of certainly unpredictable tones as indicated in the follow-up RMANOVA. First, these results indirectly replicate the findings of Bendixen et al. (2009), providing additional evidence in favor of predictive processing in audition. Moreover, the current results extend the findings of the original study by demonstrating that the underlying neural generators of the observed ERPs respond differently for different degrees of predictive certainty. Moreover, they systematically varied with the degree of predictive certainty embedded in the experimental conditions. The observed effect displays a frontocentral scalp distribution as can be seen in the topographical voltage maps of the difference between the extreme conditions (100 % minus 0 % repetition reliability; cf. Figure 2.4) which is commonly observed for ERPs elicited in primary auditory cortex. This pattern of results suggests that the underlying system flexibly adapts to different degrees of predictive certainty if upcoming auditory stimuli are potentially predictable. Following the logic of previous studies, the early onset of the effect indicates that the underlying system works in a predictable fashion (Bendixen et al., 2012). Moreover, the current pattern of results cannot easily be explained by different states of refractoriness of underlying neural populations. Since the observed ERPs were recorded in response to omissions of tones, the ERPs were not directly contaminated by the processing of actual physical stimuli and hence are assumed to better reflect intrinsic aspects of perceptual processes.

Despite some clear advantages of the applied paradigm over more traditional approaches, some questions remain unanswered. First, it cannot be ruled out that the observed effects result from differential processing of preceding stimuli. Omissions of potentially predictable tones were always preceded by frequency changes. The number of these events varied drastically between different conditions of repetition reliability (see Table 2.2) which might have altered the processing of the events in question due to different states of refractoriness of underlying neural populations. Second, the observed effects occurred immediately *after* the onset of an omission. Therefore it cannot be ruled out that the observed pattern of results reflects a process that works in a retrospective fashion.

An additional aim of the first experiment was to give further insights about possible other influences on the processing of potentially predictable tones apart from predictive certainty. In contrast to the original experiment by Bendixen et al. (2009), Experiment 1 basically allowed to differentiate between accurately predicted tones and tones which were potentially predictable but turned out to violate the prediction. Systematically analyzing different influences, like different degrees of predictive certainty vs. different states of refractoriness on early vs. later processing stages would further help to inform theories about possible post-stimulus correlates of predictive processing in audition, like e.g. the P50 component or the RP. The interaction of Stimulus Type by Window by Repetition Reliability failed to reach significance. This result indicates that there was no differential pattern of results between early and late effects of Repetition Reliability depending on the Stimulus Type (change occurred vs. repetition occurred). Significant interactions were observed for Stimulus Type by Window, Stimulus Type by Repetition Reliability and Window by Repetition Reliability, respectively. Follow-up tests of the Stimulus Type by Window interaction revealed that tones constituting a frequency change were generally more positive in the late time window as opposed to frequency repetitions irrespective of the Repetition Reliability. No such difference was observed in the early time window. At first glance, this effect might seem counter-intuitive because studies investigating the RP commonly suggest prolonged positive deflection in response to frequency repetitions. However, the prolonged positive deflections, associated with the RP, might start to unfold on later stages of processing since it initially was observed between 50 and 250 ms with slower rates of stimulation (e.g Haenschel et al., 2005). Hence the current paradigm might be unfit to validly assess the RP due to the rapid stimulation used in the experiment. Instead the current effect might be explained by a stronger neural response to frequency changes as opposed to frequency repetitions. This is commonly observed for the P50 component which is known to be suppressed in response to frequency repetitions (e.g Jerger, Biggins, & Fein, 1992). Follow-up tests of the Stimulus Type by Repetition Reliability revealed a significant effect of Repetition Reliability when averaged across both windows for frequency repetitions but not for frequency changes. This indicates that prediction-related effects might be overwritten by violations of the prediction on early stages of processing. For the frequency repetitions, the degree of Repetition Reliability was associated with more positive deflections throughout the whole post-stimulus window which might demonstrate early prediction-related contributions to the RP. Finally, a follow-up test of the Window by Repetition Reliability interaction revealed a graded linear effect of Repetition Reliability for ERPs averaged across Stimulus Type. This effect was present in the early and the late time window. Due to the absence of a three-way interaction of Stimulus Type by Window by Repetition Reliability, no clear distinctions can be made about early vs. late stages of processing between the different stimulus types depending on the predictive certainty. Even though the aforementioned effect was stronger in the early vs. the late window which suggests that prediction-related effects might have a stronger impact on early stages of processing whereas ERP effects in later stages might be driven stronger by the stimulus related processing, Experiment 1 does not provide clear evidence about differential activity for frequency changes and frequency repetitions in early versus late stages of processing, mediated by different degrees of predictive certainty.

## 2.2 Experiment 2<br>Accuracy of auditory predictions

The second experiment was conducted at the Neuropsychology Lab of the Department of Psychology (University of Oldenburg). It was aimed at systematically investigating the influence of repetition accuracy on the processing of omissions of potentially predictable vs. unpredictable tones using ERPs. The experimental logic was in line with Experiment 1 but instead of varying the conditional probability of frequency repetitions, the accuracy of auditory predictive relations was varied by introducing different degrees of deviations from perfect frequency repetition (i.e., the actual frequency of the “repeated” tone was off by a certain amount; these events will be referred to as pseudo-repetitions). Again, these levels of accuracy systematically varied in magnitude across five conditions from perceptually very inaccurate to precisely accurate.

However, there were some fundamental differences in the experimental manipulations between the first and the second experiment. In contrast to the first experiment, in the second experiment all stimuli were presented in "pairs" (i.e., every other tone was a pseudo-repetition of the previous tone [repetition possible], and a frequency change was enforced after each of these pairs [change certain]). Therefore, the distinction between potentially predictable tones following a frequency change vs. potentially predictable tones following a frequency repetition was not as well defined as in the first experiment. Hence for Experiment 2, only effects of ERPs in response to tone omissions are reported in the current Section.

As for the first experiment, an effect of repetition accuracy was expected between the extreme conditions on ERPs in response to omissions. If the experiment additionally yields a gradual modulation of this effect mediated by the different degrees of repetition accuracy, this would provide further evidence that predictive processing in audition might be fault tolerant. In the best case it would furthermore provide some information about the extent of fault tolerance and thus, would further help to characterize the proposed underlying mechanism.

### 2.2.1 Methods

**Subjects** Like in Experiment 1, twenty healthy subjects (13 female, 20-33 years old, mean age: 25.9 years) participated in the experiment. Prior to the experiment, subjects gave written consent in accordance with the Declaration of Helsinki (World Medical Association, 2013) after being informed about the nature of the experiment. Experimental procedures of Experiment 2 were approved by the local ethics committee. Subjects received course credit or modest financial compensation for their participation. The average proportion of artifact-free data was 98.24 % (standard deviation (SD): 2.49, minimum: 89.13). Data from all subjects was used for further analysis. Nineteen of the twenty subjects were right-handed (mean laterality index: 92.55) and one was left-handed (laterality index: -70) according to a German version of the Edinburgh Handedness Inventory (Oldfield, 1971).

**Experimental procedures** Subjects were seated in a comfortable chair inside an acoustically attenuated testing chamber (self-made). Isochronous tone sequences were presented binaurally via loudspeakers (Cambridge Audio S30 amplified by a Denon PMA 510 AE) positioned bilaterally at a distance of 1.5 m from the subject with a level of 70 dB [SPL]). In line with Experiment 1, participants were instructed not to pay attention to the tones while watching a self-selected, silenced movie with subtitles on a screen positioned outside of the testing chamber, visible through a glass pane. Tones were synthesized with Matlab R2011b (The MathWorks Inc., Natick, USA) and presented using the Psychophysics Toolbox extension for Matlab (Brainard, 1997). Apart from the condition-specific manipulations, the procedures were identical to Experiment 1 (tone duration: 50 ms (5 ms half-raised cosine ramps); SOA: 150 ms; frequency range: 400 - 1000 Hz; at least one semitone (5.9 %) between pairs).

**Experimental paradigm** In contrast to Experiment 1, all stimuli were presented in "pairs" (i.e., every other tone was a pseudo-repetition of the previous tone [repetition possible], and a frequency change was enforced after each of these pairs [change certain]) as in condition 5 of Experiment 1. However, here the repetition accuracy was manipulated in five conditions, again creating different levels of predictability. Frequency variation of the pseudo-repetitions was introduced ranging from 60 % to 0 % in 15 % steps relative to the whole frequency spectrum

(400 - 1000 Hz). The frequency of the pseudo-repetitions was pseudo-randomly selected within that range centered around the frequency of the previous tone, as illustrated in Figure 2.7.

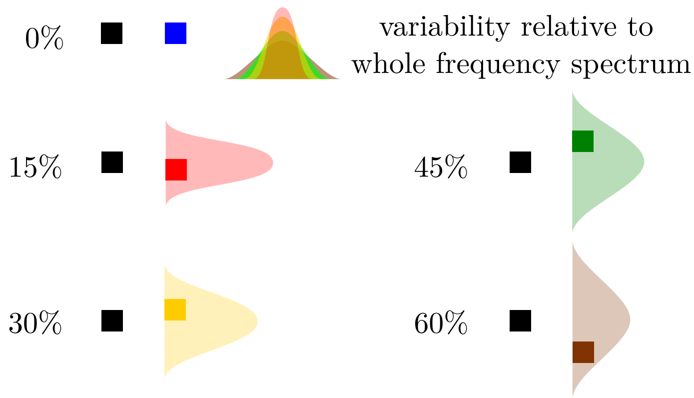


Figure 2.7: Illustration of the frequency variation range of pseudo-repetitions. The frequency of pseudo-repetitions was pseudo-randomly selected within that range, centered around the frequency of the previous tone. The randomization followed a Gaussian-like beta distribution with the parameters $\alpha = \beta = 4$.

The randomization followed a Gaussian-like beta distribution with the parameters $\alpha = \beta = 4$. In the following, this manipulation will be referred to as repetition accuracy, which is defined as the complement of the frequency pseudo-repetition inaccuracy (repetition accuracy = 1 − variation range of frequency pseudo-repetition). The randomization was furthermore constrained such that only tones falling into the overall spectrum of 400 Hz to 1000 Hz were presented (i.e., random choice of the tone frequency was repeated until a tone from the 400-1000 Hz range resulted). The choice of parameters for the frequency variation range, used in the different conditions, was based on a pilot experiment in which subjects were asked to rate whether they perceived the stimuli as pairs or rather as single tones with random frequency. Because tones were perceived as single events (rather than pairs) to the same extent from 0 % to 40 % accuracy, 40 % was chosen as the lowest value. An illustration of the experimental paradigm can be seen in Figure 2.8. In 8 % of the cases, tones were replaced with a 50 ms gap (silence). Those omissions were presented at random positions with the restriction that two omissions were at least 1050 ms apart. Four percent of the omissions were presented at positions of certainly unpredictable tones and the other 4 % of omissions were presented at positions of potentially predictable tones. In all conditions, 6000 stimuli were presented of which 240 were omissions of certainly unpredictable tones and 240

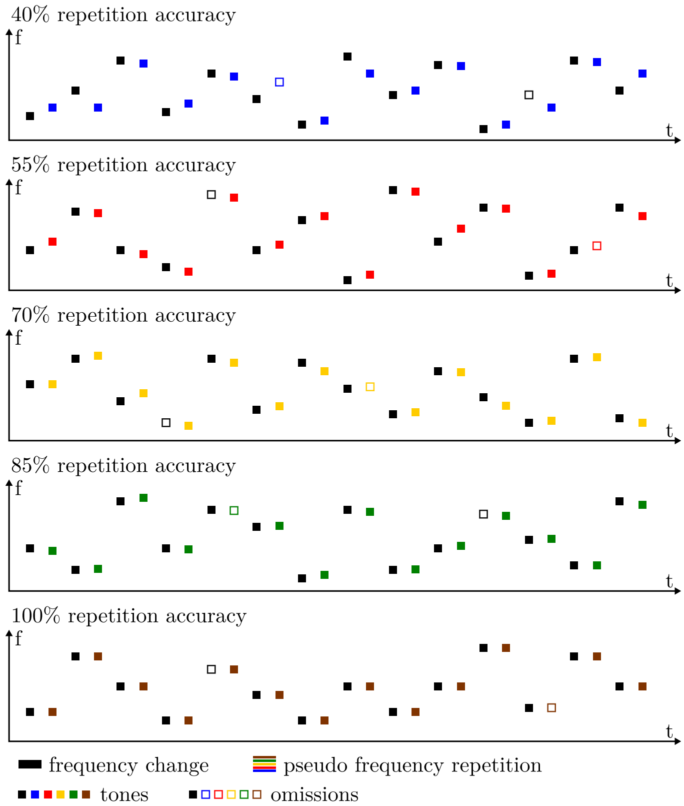


Figure 2.8: Design of Experiment 2: manipulating the repetition accuracy of predictive relations between successive tones. Every odd tone is a frequency change per definition. Every even tone is a pseudo-repetition with a variable repetition accuracy. The repetition accuracy is varied in five conditions (40 %, 55 %, 70 %, 85 % and 100 %). Eight percent of tones were replaced by omissions (equally distributed across certainly unpredictable and potentially predictable tones). Omissions were presented in random order.

were omissions of potentially predictable tones. The whole experiment consisted of 15 blocks with 2000 trials each (3 blocks per condition, presented consecutively). Condition order was counterbalanced between subjects. Net experimenting time was approximately 75 minutes. Together with electrode application and removal as well as breaks between the experimental blocks, the overall duration of the experiment amounted to 4 hours.

**Electrophysiological data acquisition** EEG was measured using a BrainAmp amplifier system (Brainproducts, Gilching, Germany) with passive Ag/AgCl electrodes from 96 scalp positions using an electrode cap with an equidistant electrode layout (Easycap, Herrsching, Germany). The horizontal EOG was measured with electrodes placed at the outer canthi of the left and right eye. The vertical EOG was obtained from separate electrodes placed below the left and right eye and from two electrodes above the eyes that were inserted in the electrode cap. The reference electrode was placed at the tip of the nose. EEG and EOG signals were amplified and recorded with a sampling rate of 500 Hz. An online lowpass filter of 249 Hz was applied to the raw data to avoid aliasing (Nyquist, 1928).

**Electrophysiological data analysis** EEG data were analyzed offline using the EEGLAB toolbox for Matlab (Delorme & Makeig, 2004). In contrast to Experiment 1, the sampling rate has been kept at its original value (500 Hz). Apart from that, the whole preprocessing procedure was identical to Experiment 1, including ICA training, ICA artifact correction, Filtering, channel interpolation and epoching. Epochs with EEG or EOG changes exceeding 100 µV were rejected from further analysis leading to an average of 1.76 % data loss. In line with Experiment 1, ERPs were baseline corrected using the pre-stimulus time window as baseline window (-150 ms to 0 ms relative to stimulus-onset). For averaging of the tone ERPs all omissions were excluded. Furthermore, tones following an omission within 600 ms of omission-onset were also excluded. That led to the exclusion of 1920 tones per condition. Grand-average ERPs were computed for all tone and omission types separately for each condition. Statistical analyses were carried out on ERPs obtained from electrode position E01 (central midline electrode placed above the vertex) which is identical to position Cz in the 10-20 system.

To investigate ERP responses to omissions of certainly unpredictable vs. potentially predictable tones across different degrees of repetition accuracy, the statistical analyses were carried out in line with Experiment 1. A within-subject RMANOVA was conducted with the factor Stimulus Type (2 levels: change certain, repetition possible) and the factor Repetition Accuracy (5 levels: 40 %, 55 %, 70 %, 85 % and 100 %) for omissions measured at Cz in the interval of 0 - 50 ms relative to stimulus-onset.

To study the topographical distribution of the ERP effects across conditions, scalp potential maps of the ERPs in the extreme conditions of each experiment were created in the respective analysis interval. For all RMANOVAs, the Greenhouse-Geisser correction was applied and the epsilon correction factor is reported whenever the sphericity assumption was violated as indicated by Mauchly's test.

### 2.2.2 Results

Grand-average ERPs of the omissions are shown in figure 2.9. Scalp topographies of the omission ERP modulation by repetition accuracy are shown in Figure 2.10. The 2 x 5 RMANOVA with the factors Stimulus Type and Repetition Accuracy for tone omissions yielded a main effect of Stimulus Type [$F(1,19)=89.190$, $p<0.00001$, $\eta^2=0.824$] indicating more positive amplitudes for omissions of potentially predictable tones in general. There was no main effect of Repetition Accuracy [$F(4,76)=0.613$, $p=0.65427$, $\eta^2=0.031$] and no Stimulus Type by Repetition Accuracy interaction [$F(4,76)=0.911$, $p=0.46206$, $\eta^2=0.046$]. Hence, no effect of Repetition Accuracy on the processing of tone omissions could be found.

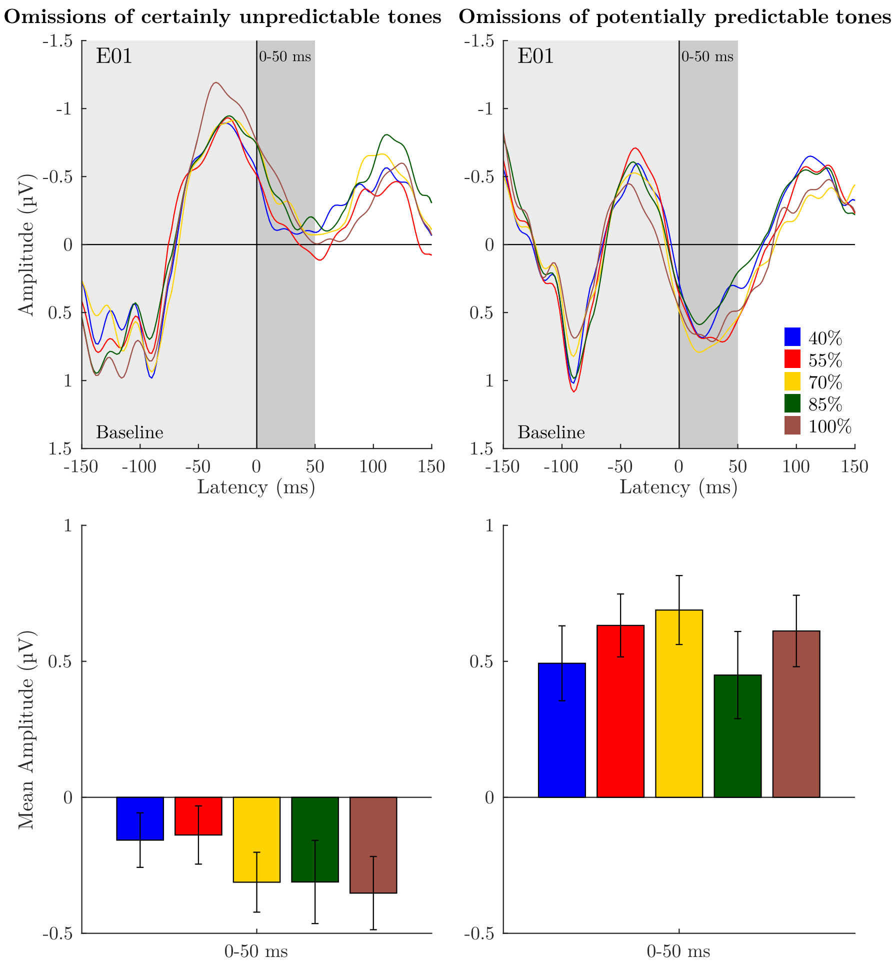


Figure 2.9: Electrophysiological Results. Upper panel: grand-average ERPs of omissions (certainly unpredictable vs. potentially predictable) across all levels of Repetition Accuracy. Lower panel: bar diagrams with mean amplitude values for each condition in the interval of 0 ms to 50 ms relative to stimulus-onset. Error bars indicate standard errors of the mean.

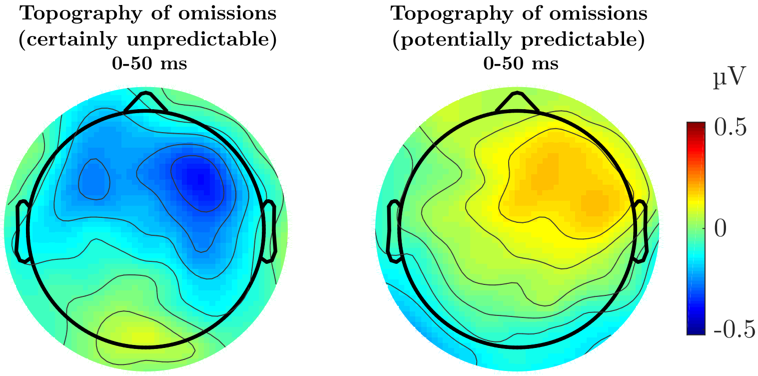


Figure 2.10: Topographical scalp voltage distribution of omission ERP modulation by repetition accuracy in the interval of 0 ms to 50 ms relative to stimulus-onset. Plotted are the difference waves of the extreme conditions (100 % minus 40 %).

### 2.2.3 Discussion

The aim of Experiment 2 was to systematically investigate the influence of repetition accuracy on ERP-correlates of omissions of certainly unpredictable vs. potentially predictable tones. Predictive relations between subsequently presented tones varied systematically by manipulating the repetition accuracy in five conditions from perceptually unpredictable (40 %) to predictable (100 %) in 15 % steps. The experiment was designed to extend the results of Experiment 1 by investigating whether the proposed underlying predictive mechanism of auditory perception forms predictions about upcoming tones even though these tones turn out to be not fully correct. If the underlying mechanism is to a certain degree fault tolerant, graded effects of repetition accuracy were expected for ERPs in response to omissions of potentially predictable tones.

First, a main effect of Stimulus Type was observed, suggesting more positive deflections of ERPs in response to omissions of potentially predictable tones as opposed to omissions of certainly unpredictable tones. This effect might be the result of the differences in the processing of preceding stimuli. Potentially predictable tones were always preceded by a frequency change, whereas certainly unpredictable tones were preceded by pseudo-repetitions. Hence, omissions of certainly unpredictable tones were more likely to be preceded by a pair of tones with equal frequency which might have resulted in different states of refractoriness of neural populations sensitive to the respective frequencies of the tones. However, this is a rather speculative explanation since pre-stimulus history was not as well defined as in Experiment 1 in which every potentially predictable tone was preceded by a frequency change and every certainly unpredictable tone was preceded by a frequency repetition. Furthermore, and in conflict with the first experiment, no effects of Repetition Accuracy were observed in the current analysis: neither for the certainly unpredictable case nor for the potentially predictable case. These contradictory results challenge the conclusions drawn from the results of Experiment 1. In the worst case, the observed effects might have been incidental findings. There is however an alternative explanation: the effects observed in Experiment 1 were already relatively small which might be explained by an insufficient SNR due to a comparatively small number of omissions presented in the experiment. Many investigations rely on such a small number of experimentally crucial stimuli since researchers are often interested in the processing of rule-violating or unexpected stimuli. However, different neural generators and different signals are likely to be involved in predictive processing, as explained in Section 1.2. A majority of previous studies focused on correlates associated with prediction error signals. However, the current investigation specifically targeted electrophysiological correlates of "purely" prediction-related activity. These signals might substantially differ in signal strength and spatial orientation (Wacongne et al., 2012) and it might be difficult to isolate them using traditional approaches.

In conclusion, two highly similar experiments were performed in which different aspects of predictive certainty were manipulated and similar effects of either Repetition Reliability or Repetition Accuracy were expected for ERPs in response to omissions of potentially predictable tones. Results of Experiment 1 support the notion of the proposed predictive mechanism which flexibly adapts to different degrees of predictive certainty embedded in the sensory context. In the second experiment, no prediction-related effects were observed at all. But even if such effects would have been present in both experiments, they cannot unequivocally be taken as evidence of prediction in a literal sense, as has been mentioned several times before, because they are usually measured *after* the onset of a potentially predictable event. In the following section, a novel approach will be presented which tries to overcome the constraints of traditional paradigms by systematically investigating ERPs shortly *before* the onset of predictable vs. unpredictable events. In Part 4, issues concerning insufficient signal strength and low statistical power will be covered in detail by accumulating data across several experiments in order to increase statistical power and to improve the SNR of possible underlying neural signals.

# 3 Pre-stimulus correlates of auditory prediction

As we have seen so far, established paradigms can help to tap into auditory processing and provide information about the underlying mechanisms which might be explained by predictive processing. A large corpus of research exists providing evidence for, and explaining many aspects of predictive processing in audition. However, the crucial problem persists that aspects of predictive processes cannot be unequivocally inferred from post-stimulus measures. No matter what, one could always argue in favor of the retrospective account.

According to different models of predictive processing (Kanai et al., 2015), there are different neural subprocesses like bottom up flow of sensory information or of prediction error signals to higher cortical layers. But according to these models there should also be a top-down flow of information, conveying the actual prediction signal. Such signals should in theory be present before the onset of an event that is expected by the system. This could already be shown in biologically plausible simulations of predictive coding in audition (Wacongne et al., 2012). If the auditory system works in accordance to such models, prediction-related neural activity should be present before the onset of predictable events. As a consequence, demonstrating graded ERP effects mediated by different degrees of predictive certainty shortly before the onset of an expected event would provide compelling evidence that the brain engages in predictive processing in a literal sense. However, note that the absence of such effects would not be evidence against such a mechanism. The neural generators of the prediction might simply not be accessible by means of EEG due to the spatial orientation of the sources or insufficient signal strength.

In this Section of the current thesis, data from the first two experiments was reanalyzed with a focus on pre-stimulus ERPs. Two more experiments are introduced to rule out alternative explanations and to further characterize prediction-related pre-stimulus ERPs.

## 3.1 Pre-stimulus effects in Experiment 1 and 2

In order to investigate if signs of auditory predictive processing can be observed in the ERPs before stimulus-onset, the data of Experiment 1 and Experiment 2 was reanalyzed with a focus on gradual effects mediated by the manipulations of predictive certainty throughout the whole epoch window. However, due to a lack of knowledge about the underlying mechanism, no expectations exist about the temporal dynamics of such a mechanism (i.e. at what time relative to the onset of a predictable event, first signs of predictive processing could possibly be observed). Moreover, the common practice of using the pre-stimulus ERP interval for baseline correction precludes observing any effects of predictability before stimulus-onset (and in the worst case, it carries pre-stimulus effects over into post-stimulus latency ranges). Hence, it was necessary to depart from certain techniques like baseline correction in order not to distort the results. For both Experiments, graded ERP effects mediated by the respective manipulation of predictive certainty (repetition reliability in Experiment 1 and repetition accuracy in Experiment 2) were expected to occur shortly before stimulus-onset of a potentially predictable event.

### 3.1.1 Methods

To investigate the influence of predictive certainty on the processing of the tones and to probe for possible pre-stimulus ERP correlates of prediction, the same preprocessing routines were used as described earlier (see Section 2.1.1 for Experiment 1 and Section 2.2.1 for Experiment 2) with the exception that no baseline correction was performed to avoid carrying over any effects from pre-stimulus to post-stimulus time windows or vice versa, thus allowing for a neutral assessment of tone processing both, before and after the onset of a stimulus. Statistical testing was performed both, for certainly unpredictable and potentially predictable tones to get a complete picture of the results. However, it should be noted that only the potentially predictable tones are informative regarding electrophysiological correlates of prediction formation in the brain. Note also that in the first Repetition Reliability condition of Experiment 1, there were only certainly unpredictable tones by definition. However, these events were used as the 0 % level of both the certainly unpredictable tones and the potentially predictable tones. Statistical analyses were carried out on ERPs obtained from a central midline electrode placed above the vertex (Cz in Experiment 1, E01 in Experiment 2).

The rationale of the statistical analysis was based on finding graded ERP effects of predictability across the five different conditions separately within each experiment. This was done by means of linear trend tests as part of within-subject RMANOVAs. Hence, in order to test whether repetition reliability influenced tone processing in Experiment 1, an RMANOVA was conducted with the factor Repetition Reliability (5 levels: 0 %, 25 %, 50 %, 75 % and 100 %). In order to test whether the repetition accuracy influenced tone processing in Experiment 2, an RMANOVA with the factor Repetition Accuracy (5 levels: 40 %, 55 %, 70 %, 85 % and 100 %) was conducted. In case of significant main effects in the RMANOVA, within-subject linear contrast analyses were performed to test for linear monotonic trends in the data. Because there were no prior hypotheses regarding the time-range of the relevant effects, first, the processing dynamics throughout the entire epoch were explored. To this aim, a running linear trend test was performed separately for each sampling point and corrected for multiple comparisons using false discovery rate (Benjamini & Hochberg, 1995). Significant intervals are indicated by the colored bars at the abscissa of each ERP plot (cf. Figures 3.1 and 3.3). The bars are colored in red if the amplitude values at the respective point were positively correlated with the model coefficients from the linear trend test, and in blue if the ERP values were negatively correlated with the model coefficients. Because in both experiments, a linear modulation of ERPs was observed during a 25 ms interval immediately preceding tone onset, as identified by the point-wise running linear trend test, this time window was chosen for the confirmatory RMANOVA as described above, before and after the onset of the stimulus. For providing a fair comparison of pre- and post-stimulus effects, an equally wide window was chosen for further processing immediately after tone onset (i.e., from 0 to 25 ms). To study the topographical distribution of the ERP effects across conditions, scalp potential maps of the ERPs in the extreme conditions of each experiment were created in the pre- and post-stimulus 25 ms ranges. Hence in Experiment 1, the difference between 100 % and 0 % Repetition Reliability was used; in Experiment 2, topographies were plotted for the difference between 100 % and 40 % Repetition Accuracy. For all RMANOVAs, the Greenhouse-Geisser correction was applied and the epsilon correction factor is reported whenever the sphericity assumption was violated as indicated by Mauchly's test.

### 3.1.2 Results

**Experiment 1** The point-wise linear trend tests for the ERPs (cf. Figure 3.1) indicate that an increase in the reliability of predictive relations between successive tones was associated with a positive deflection in the ERP around the onset of the potentially predictable tone (from 25 ms before tone onset to 50 ms after tone onset). As can be seen in Figure 3.2, this effect followed a frontocentral scalp distribution. A negative correlation was observed at the beginning and at the end of the epoch of potentially predictable tones (from -150 to -120 ms as well as from 140 to 150 ms relative to tone onset).

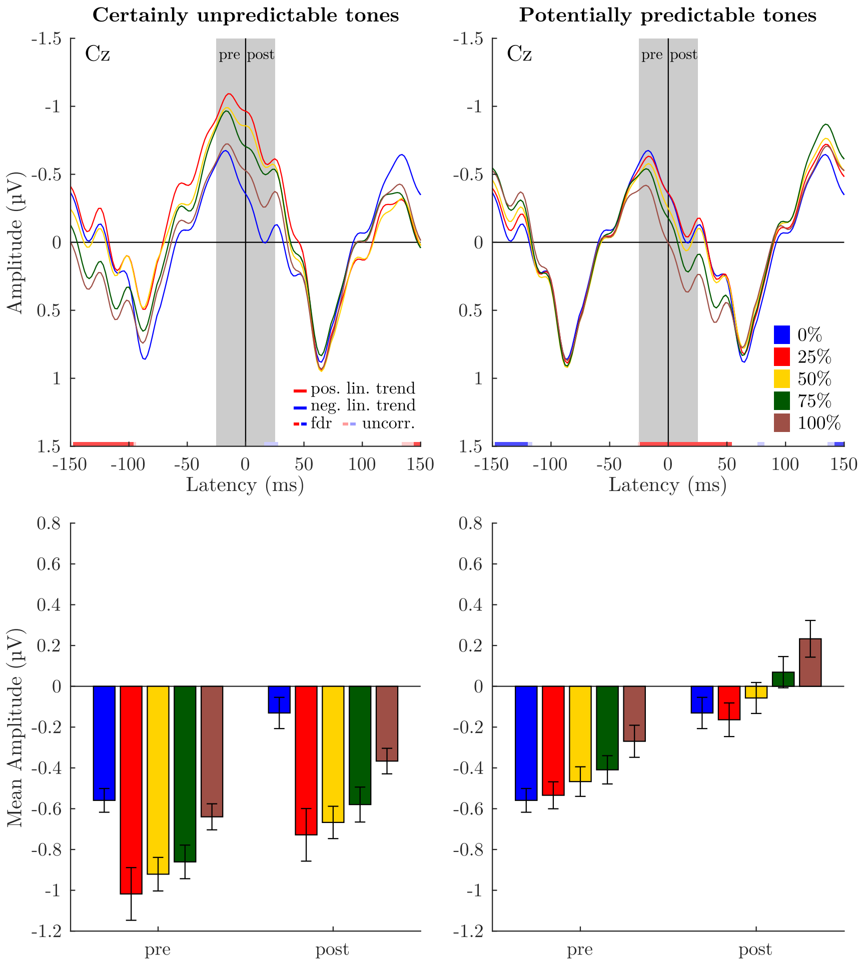


Figure 3.1: Electrophysiological Results. Upper panel: grand-average ERPs of tones (certainly unpredictable vs. potentially predictable) across all levels of Repetition Reliability. Lower panel: bar diagrams with mean amplitude values for each condition in the interval of -25 ms to 0 ms (pre-stimulus window) and 0 ms to 25 ms (post-stimulus window) relative to stimulus-onset. Error bars indicate standard errors of the mean.

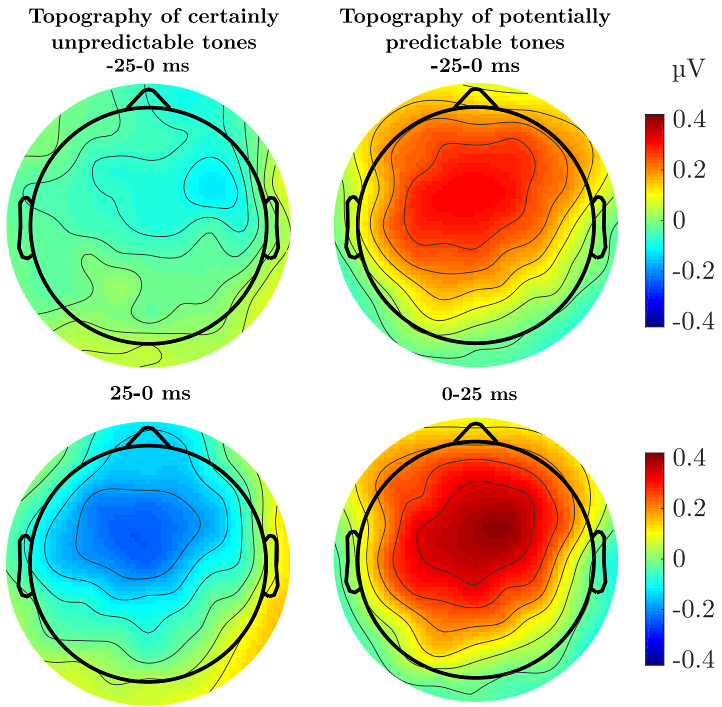


Figure 3.2: Topographical scalp voltage distribution of ERP modulation by repetition reliability in the interval of -25 ms to 0 ms (top) and 0 ms to 25 ms (bottom) relative to stimulus-onset. Plotted are the difference waves of the extreme conditions (100 % minus 0 %) for certainly unpredictable (right) and potentially predictable tones (left).

No ERP modulation by repetition reliability was observed from 120 to 25 ms before onset of the potentially predictable tone. Therefore, ERP mean amplitudes were tested within a time-range of -25 to 0 ms (pre-stimulus interval) and 0 to 25 ms (comparable post-stimulus interval) relative to tone onset. Results of this analysis are illustrated in Figure 3.1 (lower panel).

The confirmatory RMANOVA yielded a main effect of Repetition Reliability for potentially predictable tones in both windows [pre-stimulus: $F(4,76)=11.467$, $p<0.00001$, $\eta^2=0.376$; post-stimulus: $F(4,76)=17.756$, $p<0.00001$, $\eta^2=0.483$, $\epsilon=0.703$] which both followed a linear trend [pre-stimulus: $F(1,19)=28.444$, $p=0.00004$, $\eta^2=0.600$; post-stimulus: $F(1,19)=37.413$, $p<0.00001$, $\eta^2=0.663$]. This effect indicates that the ERPs shortly before the onset of potentially predictable tones varied depending on the reliability of predictive relations, with more reliable forms of predictability being associated with more positive ERP amplitudes. This effect remains stable after the onset of the tone. There were also significant main effects of Repetition Reliability for certainly unpredictable tones in the pre-stimulus window [F(4,76)=12.906, p=0.00001, $\eta^2$=0.405, $\epsilon$=0.596] and in the post-stimulus window [F(4,76)=18.242, p<0.00001, $\eta^2$=0.409, $\epsilon$=0.576] but none of these effects followed a linear trend (pre-stimulus: F(1,19)=0.001, p=0.97285, $\eta^2$<0.001; post-stimulus: F(1,19)=3.366, p=0.08227, $\eta^2$=0.150).

**Experiment 2** The point-wise linear trend test (cf. Figure 3.3) shows that the temporal dynamics of the ERP effects of potentially predictable tones closely resemble the results of tones of the same category in Experiment 1. Therefore the same window size of 25 ms was used to statistically test the ERP effects shortly before and immediately after stimulus-onset. Topographical scalp distributions of the ERP manipulations by repetition accuracy are displayed in Figure 3.4. Like in Experiment 1, there is a positively correlated linear trend prior to tone onset and lasting approximately until the end of the tone (i.e., 50 ms).

The RMANOVA with the factor Repetition Accuracy yielded a significant main effect in the pre-stimulus window [F(4,76)=6.686, p=0.00012, $\eta^2$=0.260] which followed a linear trend [F(1,19)=18.305, p=0.00041, $\eta^2$=0.491]. This effect indicates that the ERPs shortly before the onset of potentially predictable tones varied depending on the accuracy of predictive relations, with more accurate forms of predictability being associated with more positive ERP amplitudes. This effect remains stable after the onset of the tone [F(4,76)=11.750, p<0.00001, $\eta^2$=0.382], also following a linear trend [F(1,19)=36.727, p<0.00001, $\eta^2$=0.659]. As well as in Experiment 1, these effects were also present for certainly unpredictable tones [pre-stimulus: F(4,76)=14.210, p<0.00001, $\eta^2$=0.428; post-stimulus: F(4,76)=17.070, p<0.00001, $\eta^2$=0.473] but in contrast to the previous experiment, these effects also followed a linear trend [pre-stimulus: F(1,19)=28.692, p=0.00004, $\eta^2$=0.602; post-stimulus: F(1,19)=36.727, p<0.00001, $\eta^2$=0.659]. Likewise in contrast to Experiment 1, this linear trend was reversed, indicating that higher degrees of predictive accuracy are associated with more negative ERP amplitudes. It is worth noting that this effect started much earlier (at about -75 ms relative to stimulus-onset) than the positive linear trends shortly before the onset of potentially predictable tones in Experiment 1 and Experiment 2.

### 3.1.3 Discussion

The current investigation aimed at finding indicators of "true" predictions being generated in the brain – that is, to see modulations of brain responses by predictability before rather than after the onset of sensory events. Results of two experiments provide highly converging evidence that such indicators are present in electrophysiological data. Further, they shed light on the timing and polarity as well as on the flexibility of the involved brain processes. The current analysis demonstrates graded electrophysiological effects of predictive processing under different facets of uncertainty by manipulating the reliability and the accuracy of

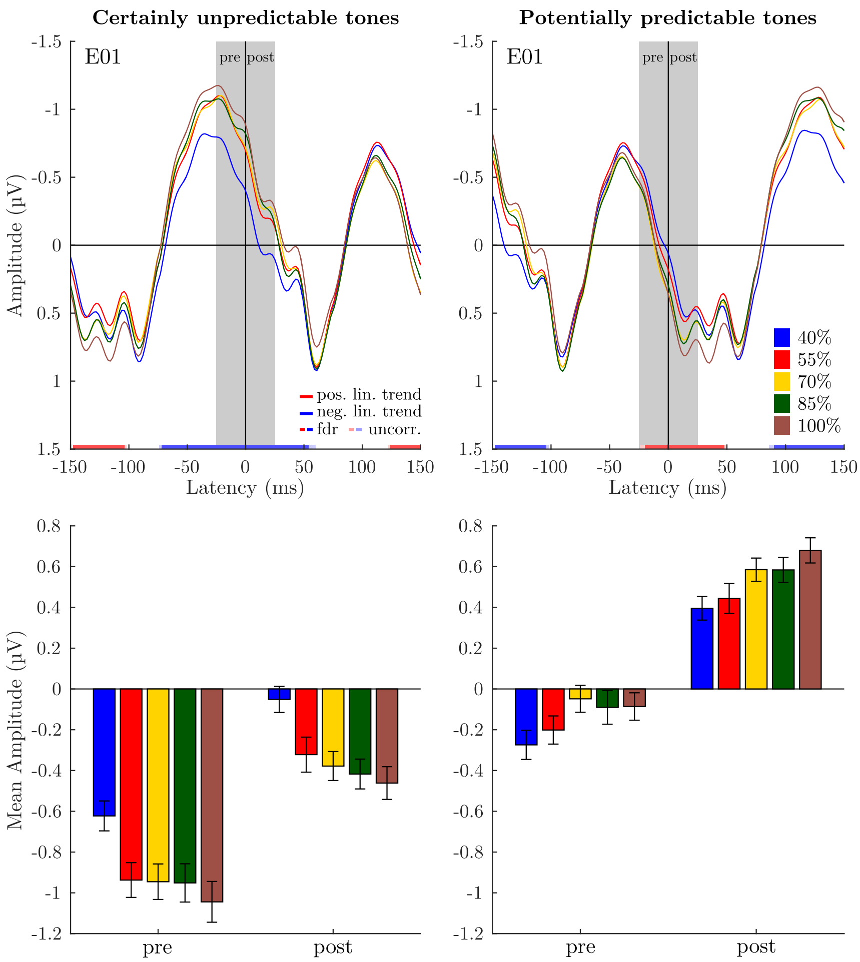


Figure 3.3: Electrophysiological Results. Upper panel: grand-average ERPs of tones (certainly unpredictable vs. potentially predictable) across all levels of Repetition Accuracy. Lower panel: bar diagrams with mean amplitude values for each condition in the interval of -25 ms to 0 ms (pre-stimulus window) and 0 ms to 25 ms (post-stimulus window) relative to stimulus-onset. Error bars indicate standard errors of the mean.

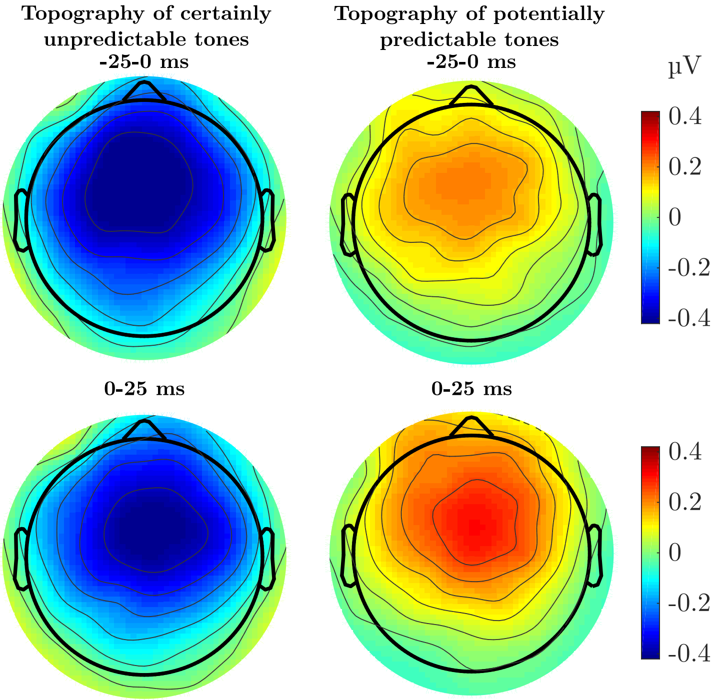


Figure 3.4: Topographical scalp voltage distribution of ERP modulation by repetition accuracy in the interval of -25 ms to 0 ms (top) and 0 ms to 25 ms (bottom) relative to stimulus-onset. Plotted are the difference waves of the extreme conditions (100 % minus 40 %) for certainly unpredictable (left) and potentially predictable tones (right).

predictive relations. Results are in line with previous studies investigating predictive processing in perception using different approaches, like match paradigms (e.g. Baldeweg, 2006; Haenschel et al., 2005), mismatch paradigms (e.g. Chennu et al., 2013), omission paradigms (e.g. Bendixen et al., 2009) or self-generation paradigms (Lange, 2011). Crucially, with this approach, signs of predictive processing prove to be present even before the onset of a stimulus, providing evidence of the predictive nature of the underlying mechanism.

As indicated by the point-wise linear trend tests, graded effects of repetition reliability (Experiment 1) and repetition accuracy (Experiment 2) became apparent approximately 25 ms prior to the onset of potentially predictable tones, lasting up until 50 ms after tone onset. In both experiments these effects followed a fronto-central scalp distribution indicating sources in primary auditory cortex. Prior to the onset of these effects, ERPs of the different conditions of predictive certainty were highly similar in both experiments. In Experiment 1 no systematic effects of Repetition Reliability were present for about 100 ms and in Experiment 2 no systematic modulation by repetition accuracy was present for about 75 ms. The absence of any ERP modulations by predictive certainty in the time ranges immediately preceding the effect-onset indicates that the observed effects are indeed related to the potentially predictable tones and are not only carry over effects from previous tones. Furthermore, the morphological similarity of the effects before and after stimulus-onset is arguably the most striking finding of the present analysis. These observations might indeed be evidence for a neural representation of top-down predictions evoked in higher auditory areas and fed back to lower sensory areas in anticipation of a stimulus as described by theoretical and computational models of predictive processing (Deneve, 2008; Rao & Ballard, 1999; Srinivasan et al., 1982). The current data provides insights about the direction of electrophysiological effects mediated by stimulus predictability. Such information is urgently needed to further understand the functionality of perceptual systems and to refine established models.

However, the current results still raise some questions. If the observed pre-stimulus effects are solely related to prediction-related parameters, they should only show up in response to potentially predictable tones. This was indeed the case for the first experiment as indicated by the point-wise linear trend tests. In Experiment 2 however, apart from the positive linear trend observed around the onset of potentially predictable tones, there was also a statistically significant negative linear trend starting at about 75 ms prior to the onset of certainly unpredictable tones. Due to this observation, carry over effects of previous tones cannot clearly be ruled out. Manipulations of predictive certainty in both experiments might have also influenced the processing of the certainly unpredictable tones which in turn might explain the condition-specific variation in the ERP deflections around the onset of the following potentially predictable tones. This issue cannot easily be resolved with the results provided by the current analysis. The experimental design used in the first two experiments doesn't allow to control for such alternative explanations. Therefore, a third experiment was performed in order to disentangle carry-over effects and prediction-related contributions of the observed ERP modulations by additionally manipulating the presentation rate. The results of this experiment will be discussed in the following sections.

## 3.2 Experiment 3 Temporal dynamics of auditory prediction

The graded pre-stimulus effects of repetition reliability and repetition accuracy observed in Section 3.1 provide evidence supporting the notion that the auditory system engages in stimulus predictions in a literal sense. However, as described in the last section, an alternative explanation might suggest that the observed effects result from the processing of the preceding tones. Due to the relatively fast presentation rate, correlates of the processing of previous tones and possible prediction-related effects shortly before the following tones might overlap. With the paradigm employed in the first two experiments, this alternative explanation cannot unequivocally be ruled out. Therefore a third experiment was designed to further clarify this issue. The experiment was carried out at the Auditory Psychophysiology Lab of the Department of Psychology (University of Oldenburg). It was designed to systematically investigate the influence of the tone presentation rate on the prediction-related effects observed in Section 3.1.2. The experimental logic was in line with Experiment 2. Again, repetition accuracy was manipulated in five conditions (40 %, 55 %, 70 %, 85 %, 100 %). Furthermore, the SOA was varied in three conditions (125 ms, 150 ms and 175 ms).

It has been demonstrated before that brain responses, associated with predictive processing, are not only affected by the information of what is likely to happen next but also by information about when it is likely to happen. Costa-Faidella, Grimm, Slabu, Díaz-Santaella, and Escera (2011) investigated the impact of timing predictability on the RP by manipulating the inter-stimulus intervals (ISI). They showed that early parts of the RP ($<$ 200 ms) are enhanced in conditions with an isochronous ISI as compared to conditions with random ISI. When predictions are formed with temporal precision, the graded pre-stimulus effects observed in Experiment 1 and 2 should also be present in Experiment 3 independent of the SOA. If however the observed effects result from the processing of the previous tones, the immediate pre-stimulus effects should only be observed in the 150 ms condition (parallel to Experiment 1 and 2). In the other two SOA conditions the onset of the effect should vary with the SOA. Strictly speaking, the effect should be shifted 25 ms towards the onset of the preceding tone (i.e. 25 ms earlier in the condition with 175 ms SOA and 25 ms later in the condition with 125 ms SOA).

### 3.2.1 Methods

**Subjects** Thirty healthy subjects participated in the experiment. Prior to the experiment, subjects were asked to give written consent in accordance with the Declaration of Helsinki (World Medical Association, 2013) after being informed about the nature of the experiment. Experimental procedures of Experiment 3 were approved by the local ethics committee. Subjects received course credit or modest financial compensation for their participation. Subjects with less than 80 % artifact-free data were rejected from further analysis, which led to the exclusion of one subject (78.14 % usable data). 27 of the remaining 29 subjects (15 female, 17-23 years old, mean age: 23.41 years) were right-handed (mean laterality index: 96.48), one subject was left-handed (laterality index: -100) and one subject was ambidextrous (laterality index: 20) according to a German version of the Edinburgh Handedness Inventory (Oldfield, 1971). The average proportion of artifact-free data for the remaining subjects was 98.44 % (SD: 1.68 %, minimum: 92.32 %).

**Experimental procedures** Subjects were seated in a comfortable chair inside an acoustically attenuated and electrically shielded testing chamber (IAC Acoustics, Niederkrüchten, Germany). Isochronous tone sequences were presented binaurally via headphones (Sennheiser HD25-1, 70Ω) with a level of 70 dB SPL. Participants were instructed not to pay attention to the tones while watching a self-selected, silenced movie with subtitles on a screen positioned outside of the testing chamber, visible through a glass pane. Tones were synthesized with Matlab R2011b (The MathWorks Inc., Natick, USA) and presented using the Psychophysics Toolbox extension for Matlab (Brainard, 1997). Apart from the condition-specific manipulations, the procedures were identical to Experiment 1 and Experiment 2 (tone duration: 50 ms (5 ms half-raised cosine ramps); frequency range: 400 - 1000 Hz; at least one semitone (5.9 %) between pairs).

**Experimental paradigm** In line with Experiment 2, all stimuli were presented in "pairs" (i.e., every other tone was a pseudo-repetition of the previous tone, and a frequency change was enforced after each of these pairs) and again, repetition accuracy was varied in five conditions by manipulating the frequency variation of the pseudo-repetitions as described in Section 2.2.1 (see also Figure 2.7). In contrast to Experiment 2, there was an additional manipulation of the SOA by shortening the original SOA by 25 ms and by prolonging it by 25 ms. There was also one condition in which the SOA remained at its original value of 150 ms. This resulted in 3 SOA conditions: 125 ms, 150 ms and 175 ms. The experimental paradigm is illustrated in Figure 3.5.

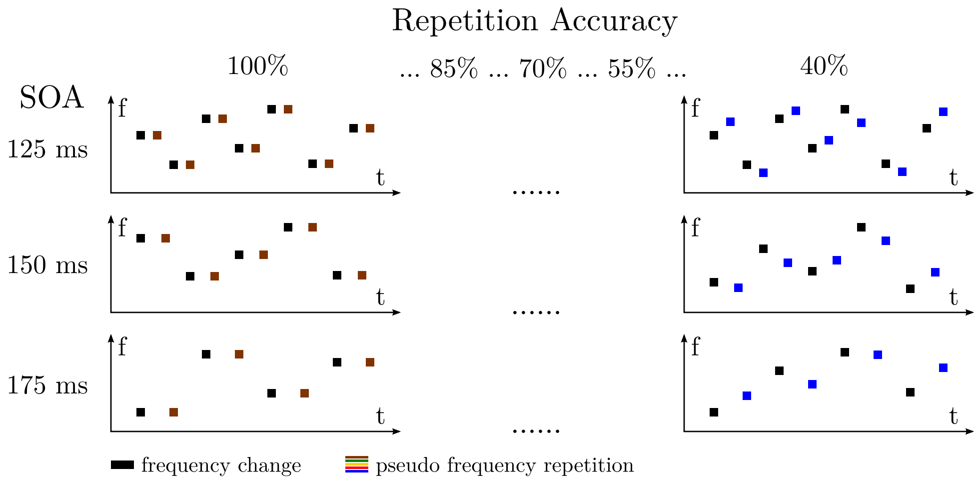


Figure 3.5: Design of Experiment 3: manipulating the repetition accuracy of predictive relations between successive tones. Every odd tone is a frequency change per definition. Every even tone is a pseudo-repetition with a variable repetition accuracy. The repetition accuracy is varied in five conditions (40 %, 55 %, 70 %, 85 % and 100 %). Each level of Repetition Accuracy was presented with 125 ms, 150 ms and 175 ms SOA in one block respectively.

Since the number of different combinations of conditions has drastically increased in this design (15 combinations: five levels of Repetition Accuracy by three levels of SOA) and since this experiment was aimed at investigating temporal dynamics of prediction-related ERP effects in anticipation of, and in response to tones, no tone-omissions were introduced in Experiment 3. Rare omissions are processed as deviant events (Yabe et al., 1997) and unavoidably lead to a decrease in usable data because immediately following tones are contaminated by the deviant response to the rare omissions and have to be excluded. For each SOA condition (1-3), five blocks were presented with a different level of Repetition Accuracy (40 % to 100 %) so that each level of Repetition Accuracy was presented with either 125 ms, 150 ms or 175 ms SOA in one block. In each block 2500 stimuli were presented. In the 125 ms SOA condition the block duration was 5.2 minutes, blocks with 150 ms SOA lasted 6.25 minutes and 175 ms SOA blocks lasted 7.29 minutes. All blocks of one SOA condition were presented consecutively and the order was counterbalanced across subjects. The order of Repetition Accuracy blocks were counterbalanced across SOA condition and subjects. Net experimenting time was approximately

94 minutes. Together with electrode application and removal as well as breaks between the experimental blocks, the overall duration of the experiment amounted to 4 hours

**Electrophysiological data acquisition** EEG was measured using a BrainAmp amplifier system (Brainproducts, Gilching, Germany) with passive Ag/AgCl electrodes from 64 scalp positions according to the 10-10 extension of the International 10-20 System (American Electroencephalographic Society, 1994) and two further electrodes at the left and right mastoid (M1, M2). The horizontal EOG was measured with electrodes placed at the outer canthi of the left and right eye. The vertical EOG was obtained from separate electrodes placed below the left and right eye and from 2 electrodes above the eyes that were inserted in the electrode cap. The reference electrode was placed at the tip of the nose. EEG and EOG signals were amplified and recorded with a sampling rate of 500 Hz. An online lowpass filter of 249 Hz was applied to the raw data to avoid aliasing (Nyquist, 1928).

**Electrophysiological data analysis** EEG data were analyzed offline using the EEGLAB toolbox for Matlab (Delorme & Makeig, 2004). The whole preprocessing procedure was identical to Experiment 2, including ICA training, ICA artifact correction, Filtering, channel interpolation and epoching (see Section 2.2.1). Epochs with EEG or EOG changes exceeding 100 µV were rejected from further analysis, leading to an average of 1.56 % data loss. In line with previous experiments, no baseline correction was performed for the assessment of peri-stimulus ERPs to avoid carrying over any effects from pre-stimulus to post-stimulus time windows or vice versa. Grand-average ERPs were computed for all tone types separately for each condition. Statistical analyses were carried out on ERPs obtained from electrode position Cz (central midline electrode placed above the vertex).

Statistical testing was performed for potentially predictable tones only since Experiment 3 aimed at investigating the temporal dynamics of the prediction-related effects, observed for potentially predictable tones in Experiment 1 and Experiment 2. The rationale of the statistical analysis was in line with Experiment 1 and 2 (for pre-stimulus effects; see Section 3.1.1). Hence, in order to test whether the repetition accuracy influenced tone processing in Experiment 3, an RMANOVA with the factor Repetition Accuracy (5 levels: 40 %, 55 %, 70 %, 85 % and 100 %)

was conducted for potentially predictable tones in all three SOA conditions for each, the pre-stimulus and the post-stimulus range used in Experiment 1 and 2 (-25 ms to 0 ms and 0 ms to 25 ms relative to stimulus-onset). In line with the previous experiments, a running linear trend test was performed separately for each sampling point and corrected for multiple comparisons using false discovery rate (Benjamini & Hochberg, 1995). Significant intervals are indicated by the colored bars at the abscissa of each ERP plot (cf. Figure 3.6). The bars are colored in red if the amplitude values at the respective point were positively correlated with the model coefficients from the linear trend test, and in blue if the ERP values were negatively correlated with the model coefficients.

To study the topographical distribution of the ERP effects across conditions, scalp potential maps of the ERPs in the extreme conditions (difference between 100 % and 40 % Repetition Accuracy) were created in the pre- and post-stimulus 25 ms ranges. For all RMANOVAs, the Greenhouse-Geisser correction was applied and the epsilon correction factor is reported whenever the sphericity assumption was violated as indicated by Mauchly's test.

### 3.2.2 Results

The point-wise linear trend tests for the ERPs (cf. Figure 3.6) indicate no linear relationships across the different levels of Repetition Accuracy in the pre- and post-stimulus ranges of all SOA conditions. The only linear trend that survived FDR correction is present around the P50 of the preceding tone in the 125 ms condition. The results of the confirmatory RMANOVAs for each window in all of the SOA conditions are illustrated in Table 3.1. These results confirm the absence of linear trends in the respective windows since none of the RMANOVAs yielded a significant effect of Repetition Accuracy which indicates that the ERPs around the onset of potentially predictable tones were not modulated by different levels of repetition accuracy. Topographical scalp distributions of the ERP manipulations by Repetition Accuracy for all levels of SOA are displayed in Figure 3.7.

Table 3.1: Statistical results. Results of the confirmatory RMANOVAs with the factor Repetition Accuracy for each window and each SOA condition.

| SOA | window | degrees of freedom | F-value | p-value | $\eta^2$ | $\epsilon$ |
|---|---|---|---|---|---|---|
| 125 ms | -25-0 ms | 4/112 | 0.191 | 0.94292 | 0.007 | 0.669 |
| 125 ms | 0-25 ms | 4/112 | 1.736 | 0.14711 | 0.058 | |
| 150 ms | -25-0 ms | 4/112 | 0.910 | 0.46051 | 0.031 | |
| 150 ms | 0-25 ms | 4/112 | 1.426 | 0.22996 | 0.048 | |
| 175 ms | -25-0 ms | 4/112 | 1.776 | 0.13862 | 0.060 | |
| 175 ms | 0-25 ms | 4/112 | 1.776 | 0.30891 | 0.042 | 0.737 |

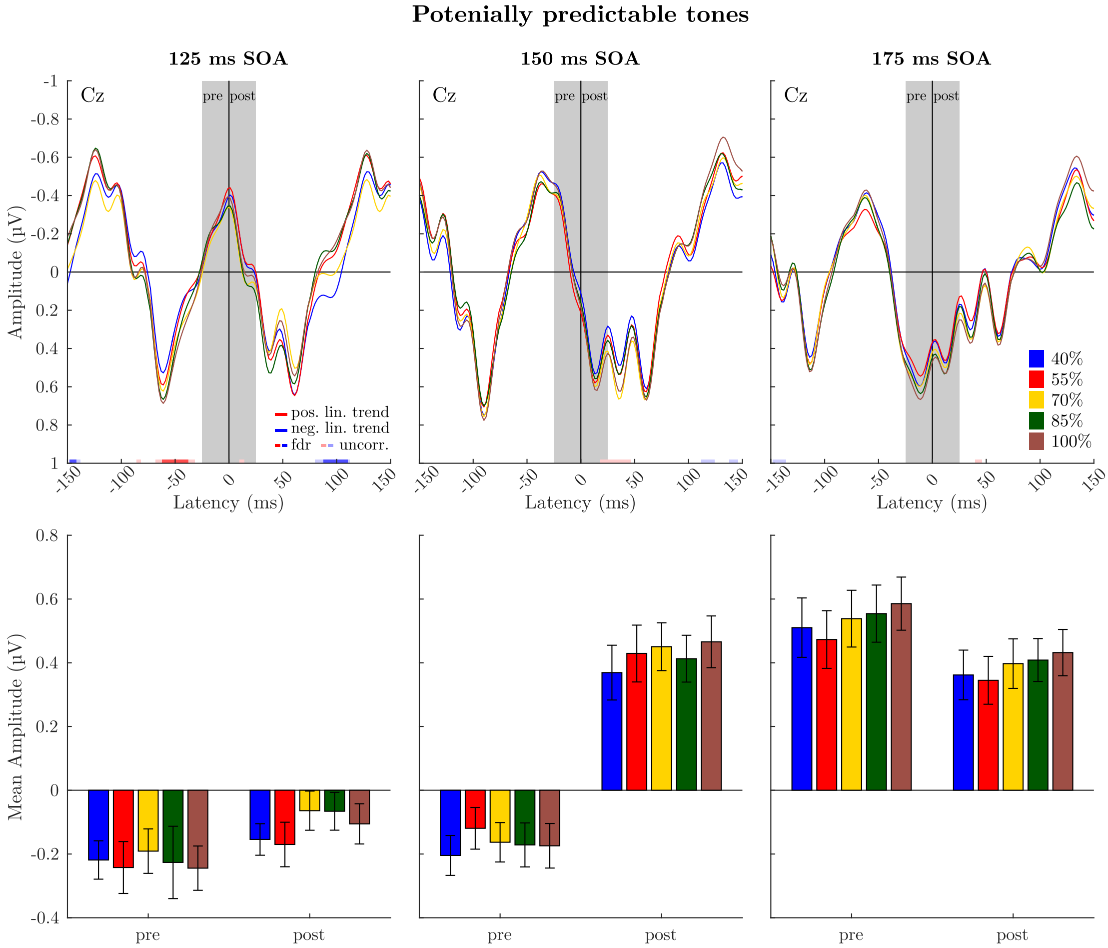


Figure 3.6: Electrophysiological Results. Upper panel: grand-average ERPs of potentially predictable tones across all levels of Repetition Accuracy for each SOA condition. Lower panel: bar diagrams with mean amplitude values for each condition in the interval of -25 ms to 0 ms (pre-stimulus window) and 0 ms to 25 ms (post-stimulus window) relative to stimulus-onset. Error bars indicate standard errors of the mean.

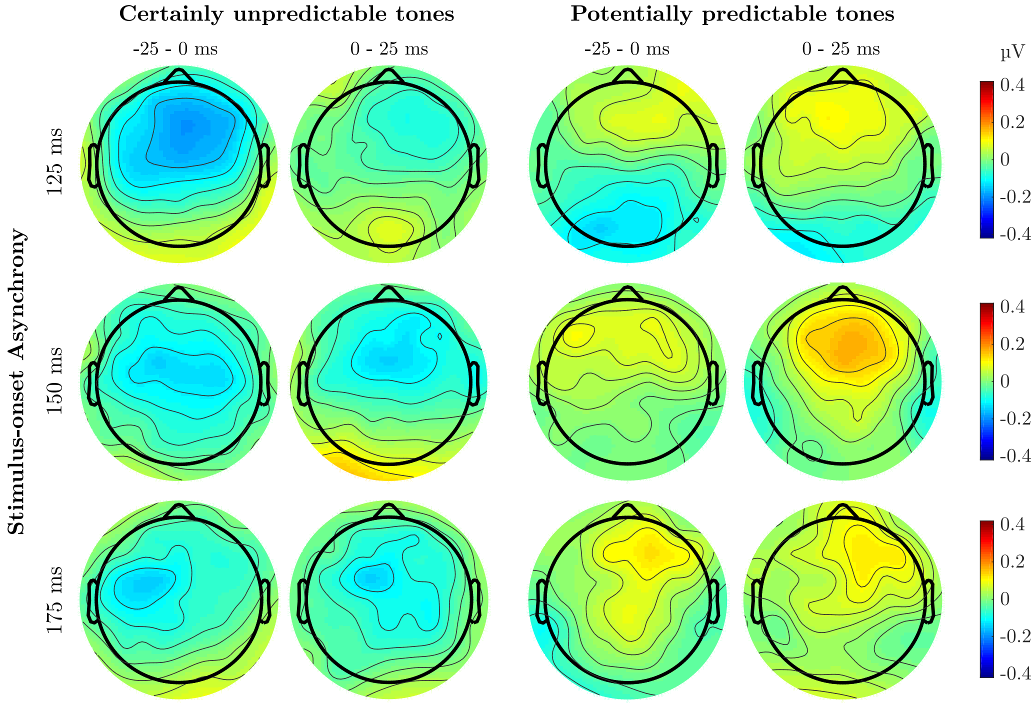


Figure 3.7: Topographical scalp voltage distribution of ERP modulation by Repetition Accuracy in the interval of -25 ms to 0 ms (1st and 3rd column) and 0 ms to 25 ms (2nd and 4th column) relative to stimulus-onset. Plotted are the difference waves of the extreme conditions (100 % minus 40 %) for certainly unpredictable (1st and 2nd column) and potentially predictable tones (3rd and 4th column) presented with different levels of SOA: 125 ms (top), 150 ms (middle) and 175 ms (bottom).

### 3.2.3 Discussion

Experiment 3 was designed to further disentangle different possible contributions to the pre-stimulus effects observed in the first two experiments. By systematically manipulating the SOA across three different conditions, the aim was to investigate the temporal dynamics of the observed pre-stimulus effects and, in the best case, to identify whether the effects are indeed driven by prediction of the upcoming tones or whether they are carry-over effects from prior stimuli. At the same time the experiment provides the opportunity for replication of the previous experiments since one condition (150 ms SOA) was basically identical to experiment 2 with the exception that there were no omissions in Experiment 3.

Contrary to the expectations, the effects observed in the first two experiments could not be replicated as indicated by point-wise linear trend tests. No gradual effects of Repetition Accuracy were observed around the onsets in all of the three

SOA conditions. Only short time ranges of significant linear trends could be observed in the 125 ms condition but in contrast to the effects observed in the first two experiments, here the effects are not clearly associated with the onset of the potentially predictable tones. There are several possible reasons that could explain the absence of any effects in Experiment 3. Given the fact that prediction-related pre-stimulus effects were observed in two independent experiments before, it seems unlikely that these effects were of incidental nature. Alternatively, the absence of any effects might be explained by an insufficient signal strength. Due to the additional manipulation of the SOA, the number of stimuli per condition had to be reduced in order to keep the experiment at a reasonable length. This might in turn have resulted in a lower SNR as compared to the first two experiments.

An alternative explanation for the absence of any effects addresses the influence of the rare tone omissions on the results observed in the first two experiments. In both the first and the second experiment, tones were occasionally replaced with silence. Such tone omissions were not included in Experiment 3 since omission-related processing was not addressed here. However, the presence of omissions at least perceptually leads to a disruption in the stimulation. This might have altered the perceptual processing of tones in general since the global structure and regularity of the sequence was affected by the presence of omissions. In Experiment 3 these omissions were not present. Hence, the tone sequences presented here were generally more continuous which might have influenced the proposed underlying predictive mechanism. For example, Abdallah and Plumbley (2009) explain this in terms of entropy of sensory input. Entropy can be explained as the quantity of different states a system can be in. A system with low entropy can therefore be seen as a highly ordered system, whereas high entropy is associated with low order. In other words: entropy can be seen as a measure of disorder. The authors describe that sensory input with low entropy (high order) is highly predictable. Therefore, the system can predict the upcoming stimulus with a high level of precision, rendering the stimulus less salient. With less entropy in the sensory input, it should theoretically be easier for the system to form predictions. In other words, the system does not have to invest as much energy in order to predict upcoming stimuli. This might have led to a weaker predictive signal in the results of Experiment 3. In the first two experiments, omissions were present and hence, the tone sequences entailed a higher degree of entropy. This might have led to stronger neural correlates of underlying predictions. To resolve this issue, a fourth experiment was designed which will be addressed in the following section.

## 3.3 Experiment 4 - The influence of omissions on stimulus processing

As discussed in the previous section, there are two options explaining the absence of any prediction-related effects in Experiment 3 regardless of the SOA. The effects observed in Experiment 1 and 2 are relatively small. Therefore a relatively high signal-to-noise ratio is necessary for the effects to prevail in the resulting ERPs. This issue will be addressed in Part 4 in greater detail. The forth and last experiment of the current thesis addresses the second option: can the missing effects be explained by the absence of omissions in Experiment 3? As explained earlier, omissions in the tone sequences might increase the salience of the signal per se and at least perceptually disrupt the stimulation, thereby generating greater entropy in the tone sequences. As a consequence, the auditory system might have to invest more energy in order to form predictions about future events. Without the omissions in the tone sequences, the stimulation becomes more regularly ordered and might in turn lose salience. The predictions formed by the system might in turn be optimized and therefore decreased in signal strength to the point where they become undetectable (at least by means of ERPs).

The last experiment was designed to clarify this issue and to investigate whether the presence of occasional omissions in the tone sequences influence the prediction-related peri-stimulus effects obtained in Experiment 1 and 2 (see Section 3.1.2) while keeping the remaining experimental parameters constant. The experiment was carried out at the Cognitive Systems Lab of the Department of Physics at Chemnitz University of Technology. The experimental logic was in line with Experiment 2. Again, repetition accuracy was manipulated in five conditions (40 %, 55 %, 70 %, 85 %, 100 %). Furthermore, the tone sequences were either presented with or without the occasional omissions. If the prediction-related pre-stimulus effects can be explained by the presence of omissions, graded effects of Repetition Accuracy should be observable shortly before stimulus-onset of potentially predictable tones. This effect should only be observed in conditions where omissions where present. Additionally, ERPs in response to tone-omissions were analyzed in accordance to Experiment 1 and 2 to further clarify the pattern of results.

### 3.3.1 Methods

**Subjects** Twenty healthy subjects (11 female, 22-34 years old, mean age: 26.3 years) participated in the experiment. Prior to the experiment, subjects gave written consent in accordance with the Declaration of Helsinki (World Medical Association, 2013) after being informed about the nature of the experiment. Experimental procedures of Experiment 4 were approved by the local ethics committee. Subjects received course credit for their participation. The average proportion of artifact-free data was 98.68 % (SD: 1.82 %, minimum: 93.04 %). Data from all subjects was used for further analysis. All subjects were right-handed (mean laterality index: 91.61) according to a German version of the Edinburgh Handedness Inventory (Oldfield, 1971).

**Experimental procedures** Subjects were seated in a comfortable chair inside an acoustically attenuated and electrically shielded testing chamber (IAC Acoustics, Niederkrüchten, Germany). Isochronous tone sequences were presented binaurally via headphones (Sennheiser HD25-1, 70Ω) with a level of 70 dB SPL. Participants were instructed not to pay attention to the tones while watching a self-selected, silenced movie with subtitles on a screen positioned outside of the testing chamber, visible through a glass pane. Tones were synthesized with Matlab R2015b (The MathWorks Inc., Natick, USA) and presented using the Psychophysics Toolbox extension for Matlab (Brainard, 1997). Apart from the condition-specific manipulations, the procedures were identical to Experiment 1 and Experiment 2 (tone duration: 50 ms (5 ms half-raised cosine ramps); SOA 150 ms; frequency range: 400 - 1000 Hz; at least one semitone (5.9 %) between pairs).

**Experimental paradigm** In line with Experiment 2, all stimuli were presented in "pairs" (i.e., every other tone was a pseudo-repetition of the previous tone, and a frequency change was enforced after each of these pairs) and again, repetition accuracy was varied in five conditions by manipulating the frequency variation of the pseudo-repetitions as described in Section 2.2.1 (see also Figure 2.7). The experimental paradigm is illustrated in Figure 3.8. In another condition, the presence of occasional omissions in the tone sequences was varied across two levels (omissions present vs. omissions absent). In blocks containing omissions, 8 % of the tones were replaced with a 50 ms gap (silence). Those omissions were presented at random positions with the restriction that two omissions were at least 1050

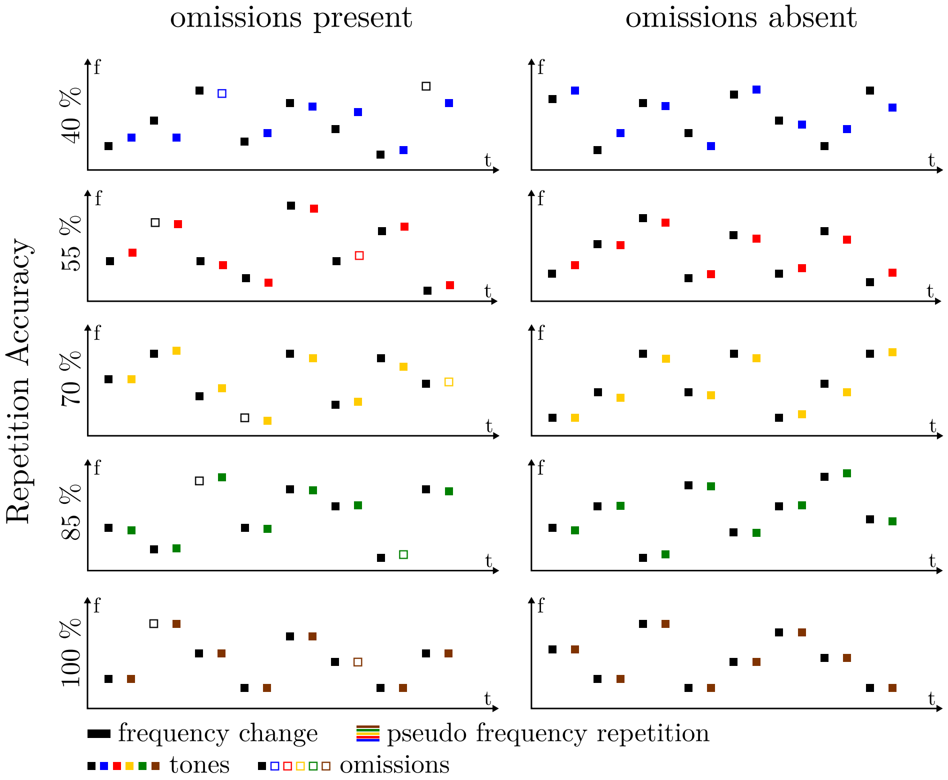


Figure 3.8: Design of Experiment 4: manipulating the repetition accuracy of predictive relations between successive tones. Every odd tone is a frequency change per definition. Every even tone is a pseudo-repetition with a variable repetition accuracy. The repetition accuracy is varied in five conditions (40 %, 55 %, 70 %, 85 % and 100 %). In half of the blocks, 8 % of tones were replaced by omissions (equally distributed across certainly unpredictable and potentially predictable tones). Omissions were presented in random order. In the other half of the blocks, no omissions were present.

ms apart. Four percent of the omissions were presented at positions of certainly unpredictable tones and the other 4 % of omissions were presented at positions of potentially predictable tones. The experiment consisted of 20 blocks with 1800 stimuli of which 10 blocks contained tone omissions (144 omissions of certainly unpredictable tones; 144 omissions of potentially predictable tones) and another 10 blocks did not contain omissions. Blocks of one condition were presented consecutively. Within each of those 10 blocks, there were two blocks of each Repetition Accuracy condition, presented consecutively. The order of Repetition Accuracy blocks was counterbalanced between the levels of the Omission condition and between subjects. The order of presentation for the levels of the Omission condition was also counterbalanced between subjects. All in all, 3600 stimuli were presented per combination of Omission (absent vs. present) by Repetition Accuracy (40 %, 55 %, 70 %, 85 %, 100 %). One block lasted 4.5 minutes. Net experimenting time was approximately 90 minutes. Together with electrode application and removal as well as breaks between the experimental blocks, the overall duration of the experiment amounted to 4 hours.

**Electrophysiological data acquisition** The same setup has been used as in Experiment 3: a BrainAmp amplifier with a 64 channel layout and further electrodes for the Mastoids and for the EOG, reference to nose, sampled at 500 Hz, 249 Hz lowpass filtered (see also Section 3.2.1).

**Electrophysiological data analysis** EEG data were analyzed offline using the EEGLAB toolbox for Matlab (Delorme & Makeig, 2004). The whole preprocessing procedure was identical to Experiment 2, including ICA training, ICA artifact correction, filtering, channel interpolation and epoching (see Section 2.2.1). Epochs with EEG or EOG changes exceeding 100 µV were rejected from further analysis leading to an average of 1.45 % data loss. For the analysis of ERPs in response to omissions, ERPs were baseline corrected in line with Experiment 1 and 2, using the pre-stimulus time range as baseline window (-150 ms to 0 ms relative to stimulus-onset). In line with previous experiments, no baseline correction was performed for the assessment of peri-stimulus ERPs of tones to avoid carrying over any effects from pre-stimulus to post-stimulus time windows or vice versa. Grand-average ERPs were computed for all tone types separately for each condition. As for the previous experiments, statistical analyses were carried out on ERPs obtained from electrode position Cz (central midline electrode placed above the vertex).

To investigate ERP responses to omissions across different degrees of repetition accuracy, the statistical analyses were carried out in line with Experiment 1 and Experiment 2. A within-subject RMANOVA was conducted with the factor Stimulus Type (2 levels: change certain, repetition possible) and the factor Repetition Accuracy (5 levels: 40 %, 55 %, 70 %, 85 % and 100 %) for omissions measured at Cz in the interval of 0 - 50 ms relative to stimulus-onset. For the assessment of possible influences of omissions on the peri-stimulus effects of Repetition Accuracy obtained in Experiment 1 and 2, and in line with Experiment 3, statistical testing was performed only for potentially predictable tones. A within-subject RMANOVA with the factors Omission (2 levels: absent vs. present) and Repetition Accuracy (5 levels: 40 %, 55 %, 70 %, 85 %, 100 %) were conducted for each, the pre-stimulus and the post-stimulus window. In line with the previous experiments, a running linear trend test was performed separately for each sampling point and corrected for multiple comparisons using false discovery rate (Benjamini & Hochberg, 1995). Significant intervals are indicated by the colored bars at the abscissa of each ERP plot (cf. Figure 3.11). The bars are colored in red if the amplitude values at the respective point were positively correlated with the model coefficients from the linear trend test, and in blue if the ERP values were negatively correlated with the model coefficients.

To study the topographical distribution of the ERP effects across conditions, scalp potential maps of the ERPs in the extreme conditions were created in the respective analysis interval. For all RMANOVAs, the Greenhouse-Geisser correction was applied and the epsilon correction factor is reported whenever the sphericity assumption was violated as indicated by Mauchly's test.

### 3.3.2 Results

**ERPs of tone omissions** Grand-average ERPs of the omissions are shown in figure 3.9. Scalp topographies of the omission ERP modulation by Repetition Accuracy are shown in Figure 3.10. The 2 x 5 RMANOVA with the factors Stimulus Type and Repetition Accuracy for tone omissions yielded a main effect of Stimulus Type $F(1,19)=12.243$, $p<0.00001$, $\eta^2=0.663$] indicating generally more positive amplitudes in the range of 0 ms to 50 ms for potentially predictable tones as compared to certainly unpredictable tones. No Interaction of Stimulus Type by Repetition Accuracy [$F(4,76)=0.434$, $p=0.78400$, $\eta^2=0.022$] and no main effect of Repetition Accuracy was found $F(4,76)=0.609$, $p=0.65742$, $\eta^2=0.031$], indicating that different degrees of repetition accuracy neither had an effect on the processing of omissions of certainly predictable tones nor on omissions of potentially predictable tones.

**ERPs of potentially predictable tones** Grand-average ERPs of tones with omissions vs. tones without omissions are shown in figure 3.11. The 2 x 5 within-subject RMANOVA for the pre-stimulus window neither yielded a main effect of Omission [$F(1,19)=2.370$, $p=0.14015$, $\eta^2=0.111$], Repetition Accuracy [$F(4,76)=1.713$, $p=0.14016$, $\eta^2=0.083$], nor an interaction of Omission by Repetition Accuracy [$F(4,76)=0.309$, $p=0.87086$, $\eta^2=0.016$].

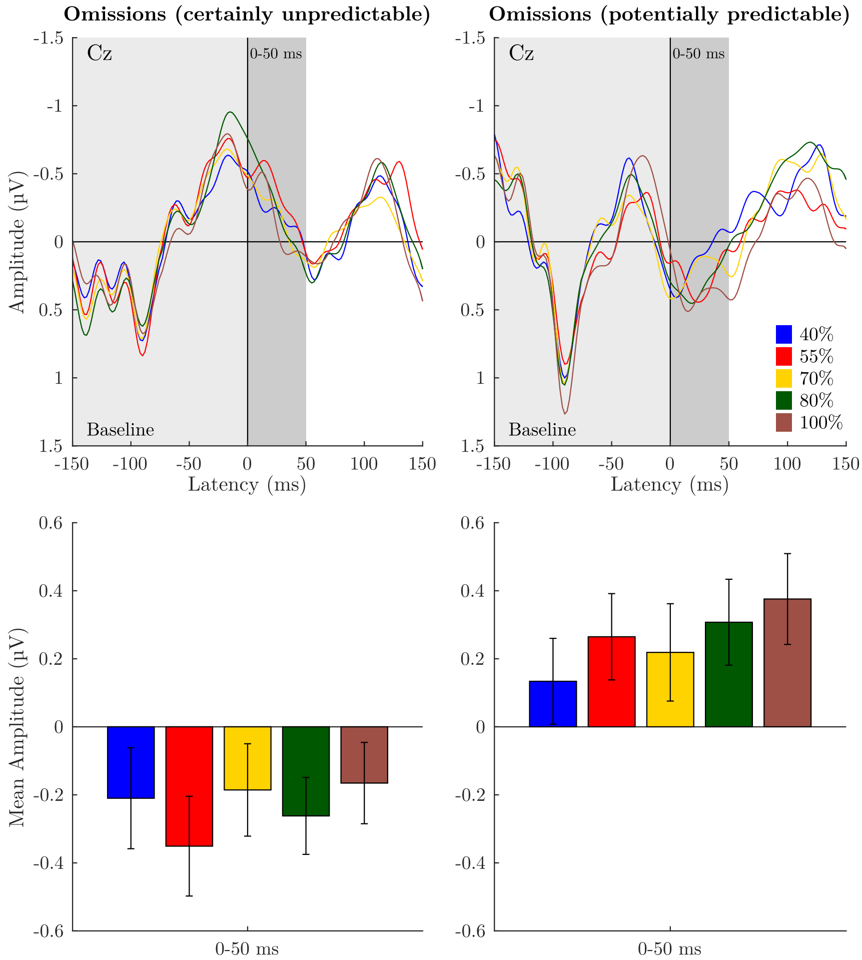


Figure 3.9: Electrophysiological Results. Upper panel: grand-average ERPs of omissions (certainly unpredictable vs. potentially predictable) across all levels of Repetition Accuracy. Lower panel: bar diagrams with mean amplitude values for each condition in the interval of 0 ms to 50 ms relative to stimulus-onset. Error bars indicate standard errors of the mean.

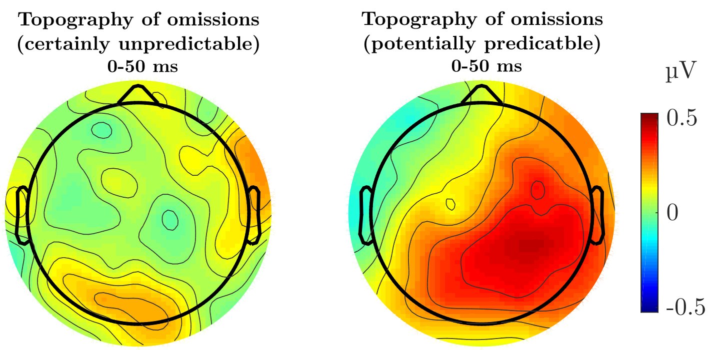


Figure 3.10: Topographical scalp voltage distribution of omission ERP modulation by Repetition Accuracy in the interval of 0 ms to 50 ms relative to stimulus-onset. Plotted are the difference waves of the extreme conditions (100 % minus 40 %).

These results indicate that there was no modulation of ERPs by different degrees of repetition accuracy in the pre-stimulus window. For the post-stimulus window, the RMANOVA yielded a significant main effect of Repetition Accuracy [F(4,76)=2.548, p=0.04605, $\eta^2$=0.118] which followed a linear trend [F(1,19)=5.614, p=0.02856, $\eta^2$=0.228]. No main effect of Omission [F(1,19)=2.766, p=0.11267, $\eta^2$=0.127] and also no interaction of Omission by Repetition Accuracy [F(4,76)=0.755, p=0.55807, $\eta^2$=0.038] could be found. The current results indicate that the ERPs of potentially predictable tones varied depending on the accuracy of predictive relations, with more accurate forms of predictability being associated with more positive ERP amplitudes. This is only the

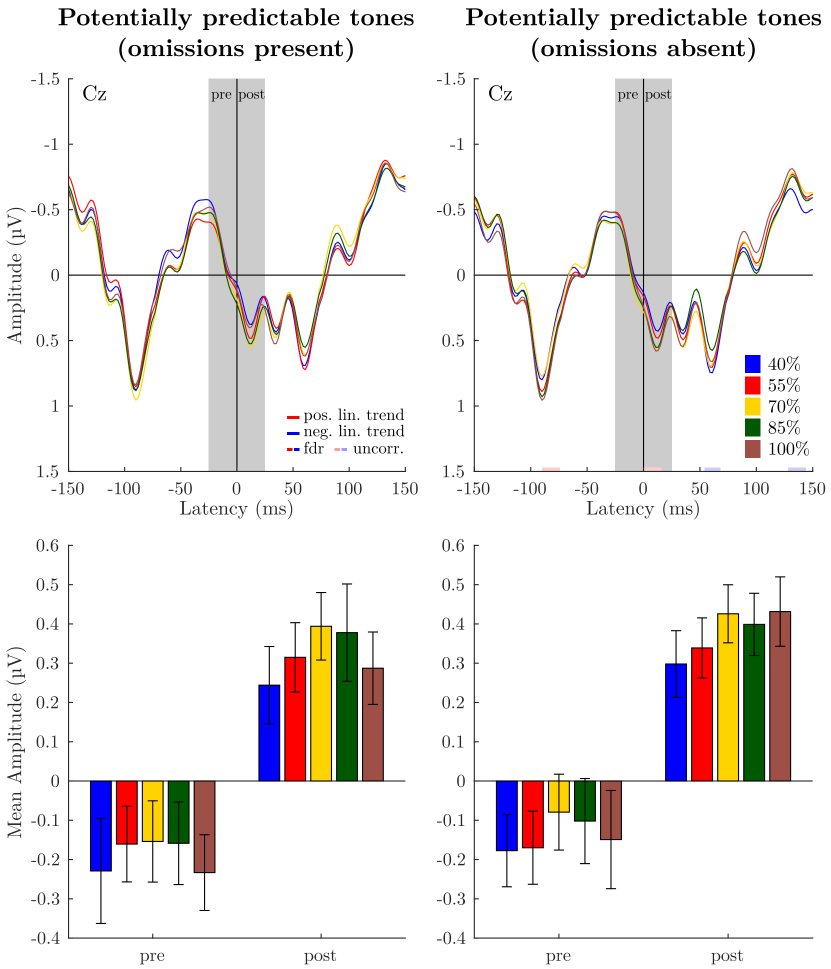


Figure 3.11: Electrophysiological Results. Upper panel: grand-average ERPs of potentially predictable tones (omission present vs. omission absent) across all levels of Repetition Accuracy. Lower panel: bar diagrams with mean amplitude values for each condition in the interval of -25 ms to 0 ms (pre-stimulus window) and 0 ms to 25 ms (post-stimulus window) relative to stimulus-onset. Error bars indicate standard errors of the mean.

case for ERPs in the range of 0 ms to 25 ms relative to stimulus-onset and irrespective of the presence of omissions in the tone sequences. Since there was no effect induced by the presence of omissions and in order to increase the SNR, ERPs were reanalyzed and both, ERPs of potentially predictable tones with and without omissions were averaged together (cf. Figure 3.12).

A within-subject RMANOVA with the factor Repetition Accuracy was conducted for each, the pre-stimulus and the post-stimulus window. In the pre-stimulus window, no effect of Repetition Accuracy was found [F(4,76)=1.780, p=0.14152, $\eta^2$=0.086]. There was a marginally significant main effect of Repetition Accuracy in the post-stimulus window [F(4,76)=2.747, p=0.03425, $\eta^2$=0.126] which followed a linear trend [F(1,19)=6.574, p=0.01899, $\eta^2$=0.257]. However, this effect does not survive Bonferroni correction. Furthermore, this pattern of results is reflected by a positive linear trend in the first 15-20 ms of the post-stimulus window indicated by the point-wise running linear trend test in Figure 3.12. Note, that this effect did not survive FDR correction. Scalp topographies of the Repetition Accuracy contrasts for potentially predictable tones are illustrated in Figure 3.13.

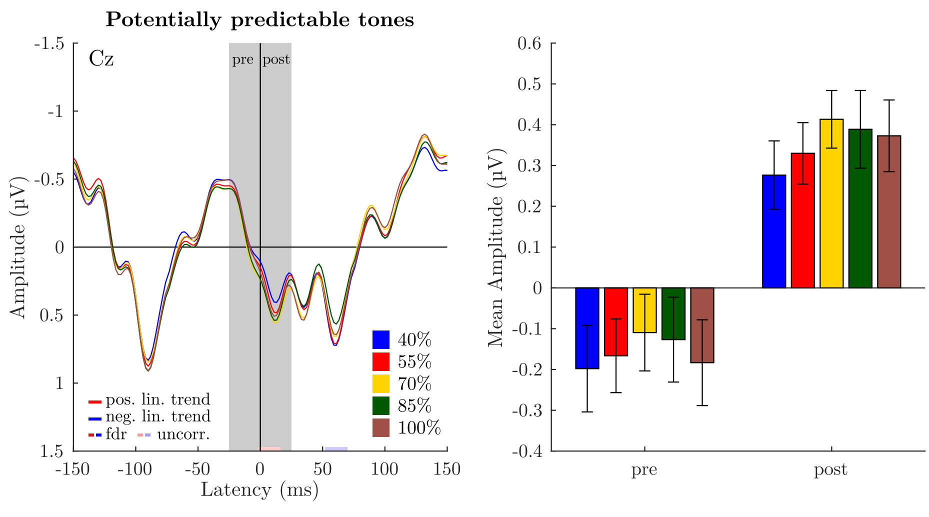


Figure 3.12: Electrophysiological Results. Left: grand-average ERPs of potentially predictable tones (averaged across omission present and omission absent) across all levels of Repetition Accuracy. Right: bar diagrams with mean amplitude values for each condition in the interval of -25 ms to 0 ms (pre-stimulus window) and 0 ms to 25 ms (post-stimulus window) relative to stimulus-onset. Error bars indicate standard errors of the mean.

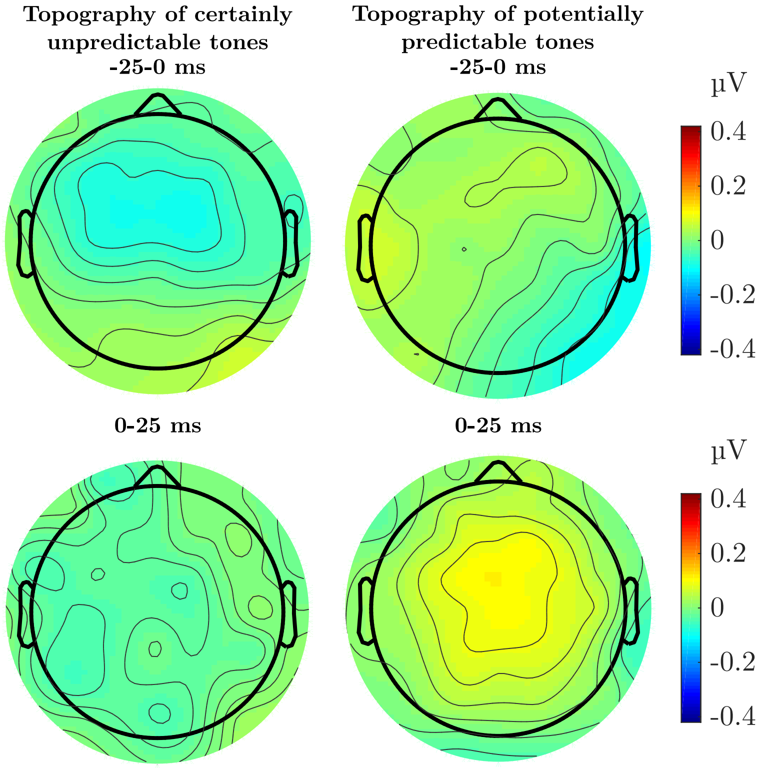


Figure 3.13: Topographical scalp voltage distribution of ERP modulation by Repetition Accuracy in the interval of -25 ms to 0 ms (top) and 0 ms to 25 ms (bottom) relative to stimulus-onset. Plotted are the difference waves of the extreme conditions (100 % minus 40 %) for certainly unpredictable (left) and potentially predictable tones (right).

### 3.3.3 Discussion

Experiment 4 was designed to investigate whether the lack of prediction-related effects in Experiment 3 can be explained by the absence of occasional omissions within the tone sequences. If the pre-stimulus effects observed in the first two experiments would again be obtained in the conditions including omissions, this would suggest that the underlying system responds differently depending on the global regularity of the tone sequence. Alternatively, if the effects would be regained independent of the presence of the omissions then the absence of effects in Experiment 3 could likely be explained by insufficient signal strength since the number of stimuli was increased in Experiment 4 in order to rule out this factor. If however the effects cannot be regained at all, this would have more severe implications regarding the validity of the results observed in Experiment 1 and 2.

The results suggest the latter case. The effects observed in Experiment 1 and 2 could not be regained irrespective of the presence of occasional omissions within the tone sequences. Additionally, the effects didn't even survive in the post-stimulus window, so even after stimulus-onset no effects of repetition accuracy could be observed. Furthermore, no effects of repetition accuracy were observed for omissions of potentially predictable tones. Considering the current state of results, it is very difficult to draw a clear conclusion. Regarding the ERPs in response to omissions of predictable tones, data from Experiment 1 provides evidence in favor of predictive processing and data from Experiment 2 and 4 are lacking any prediction-related effects. Regarding the pre-stimulus ERPs of potentially predictable tones, data of two experiments are in favor of auditory predictions, whereas 2 experiments did not show any effects associated with the manipulations of predictive certainty.

From a conservative perspective, this implies that the results obtained from Experiment 1 and 2 lack sufficient robustness because the results could not be replicated and hence should be considered as incidental findings. From a more liberal perspective, the positive results from the first two experiments should at least be interpreted carefully. Either way, the results obtained within the experiments of the current thesis cannot be considered as unequivocal evidence - neither in favor of predictive processing in audition nor against it. This puzzling situation is not an isolated incidence within modern life sciences. According to a study by Anderson et al. (2015), quite the opposite seems to be the case for psychological and neuroscientific fields of research. That is, a large number of published research is likely to be contaminated by false positive findings. In the next section this issue will be further discussed and an attempt will be made to resolve some of these issues.

# 4 Results across experiments

In recent years there has been a growing concern about what is sometimes referred to as a 'replication crisis' in many areas of psychology and neuroscience. According to a review by Barch and Yarkoni (2013), this issue refers to a high risk of producing and publishing false positive results (Masicampo & Lalande, 2012; Vul & Pashler, 2012) as a consequence of several reasons, like ubiquitous low power (Button et al., 2013), conflicts of interest (Bakker & Wicherts, 2011; Ioannidis, 2005), misaligned incentives and questionable research practice (John et al., 2012) which often results in what is commonly referred to as p-hacking (Simmons et al., 2011).

According to Simmons et al. (2011), p-hacking refers to the problem that researchers are more likely to falsely find evidence that an effect exists than to correctly find evidence that it does not (despite the nominal endorsement of a low rate of false positive findings of commonly less than 5 %). The authors discuss that a major contribution to this problem is based on the concept of *researcher degrees of freedom*. This concept describes the decisions researchers have to make during data collection and data analysis, like the amount of data that should be collected, exclusion of observations, the combination and comparison of conditions and measures and the choice of control variables. Simmons et al. (2011) provide some guidelines for researchers and reviewers to overcome these problems such as the reporting of all measured variables and all experimental conditions including failed manipulations as well as the formulation of rules for terminating data collection before data collection begins (among others).

Another major methodological problem in empirical psychological research is the practice of reporting results of chronically under-powered experiments. According to Button et al. (2013), low statistical power increases the chance of producing both, false negative and false positive results. False negative results suggest the absence of an effect that is actually present whereas false positive results erroneously reflect effects that are actually not true and are subject to sampling variations and random errors in the measurement. Furthermore, the authors argue that even if a low-powered study detects a true effect, the effect-sizes are likely to be overestimated: due to the low power of the sample, tests only reach statistical significance if a certain threshold of effect size is exceeded. The underlying effects might be much smaller in general but the test - due to insufficient statistical power - only detects effects that are of greater magnitude due to random variation in the data. Such tests inadvertently reveal so-called inflated effect sizes. Additionally, variations in the data analysis (e.g. exclusion of subjects) are more likely to influence the outcome of statistical effects in studies using smaller sample sizes as opposed to studies with greater statistical power. In other words: low-powered studies are less robust and more likely to be affected by small variations in the analysis as well as random variations in the data.

Research presenting novel and statistically significant findings is more likely to be published which creates strong incentives for researchers to quickly generate results by selectively reporting their procedures and findings as well as using small sample sizes as described above. As a consequence, a lot, if not most, of conclusions drawn from biomedical research findings might be false (Ioannidis, 2005). Such publication biases are present across all fields of experimental research, and in fact could also be demonstrated for research findings in the fields of experimental psychology and neuroscience which has recently been revealed by Anderson et al. (2015). In an extensive attempt to characterize the reproducibility of psychological research findings, the authors conducted replications of 100 studies published in peer-reviewed journals across different branches of psychological research. Of the original studies, 97 % reported significant results. However, only 36 % of the original effects could be replicated. The authors conclude that a large portion of replications did not reproduce evidence supporting the original results despite using designs with high statistical power, original materials and review in advance for methodological fidelity. Related methodological problems have just recently been revealed in the field of neuroimaging. Certain statistical techniques could be demonstrated to yield up to 70 % of false positive results instead of the theoretically assumed probability of 5 % (Eklund, Nichols, & Knutsson, 2016). Results like this fueled the already ongoing debate about the reliability of psychological and neuroscientific research which is still in process at the time of writing of this thesis. However, the results emphasize the importance of replication in experimental research and demonstrate current weaknesses and pitfalls of the scientific process in many fields of modern life sciences.

As discussed before, some of the results presented in the current thesis are partly of conflicting nature and might in turn be subject to some of the statistical and methodological pitfalls described throughout this section. Luckily, there are techniques to tackle some of the aforementioned problems like aggregation of data across several studies or rigorous replications of experiments. In an attempt to approach some of these problems, the main results presented in Part 2 and Part 3 of the current thesis were re-evaluated by aggregating data across the conducted Experiments wherever it was feasible in order to increase the statistical power and to reduce the probability of finding false positive or false negative results induced by insufficient power. The basic experimental logic applied within the scope of the current thesis was more or less identical in all of the conducted experiments: isochronous tone sequences were passively presented to subjects that were watching a self selected, silenced movie with subtitles. Predictive relations between successive stimuli were based on tone frequency in all of the experiments. Despite some differences in the nature of the experimental manipulations between the first and the remaining three experiments, there was a gradual manipulation of predictive certainty in all of the applied paradigms (five levels of repetition reliability in Experiment 1 and likewise, five levels of repetition accuracy in Experiments 2, 3 and 4). Omissions were present only in Experiment 1, 3 and 4. Hence, for the analysis of omissions, data of these three experiments was pooled together and reanalyzed using the same techniques and parameters as for each individual experiment. Scrutinizing the prediction-related pre-stimulus effects did not require the presence of omissions within the tone sequences. Therefore data of all experiments was merged in order to reanalyze pre-stimulus effects.

## 4.1 Methods

**Analysis of omissions** In order to analyze ERP responses to tone omissions across experiments, ERPs of subjects in Experiment 1, 2 and 4 were pooled together to form grand-average ERPs of omissions. Data from Experiment 3 was not included in the current analysis because no omissions were presented in this experiment. This resulted in a sample size of 60 subjects. EEG preprocessing and data analysis were carried out in line with the analysis of the respective experiment (for details see Section 2.1.1 for Experiment 1, Section 2.2.1 for Experiment 2 and Section 3.3.1 for Experiment 4). The resulting omission ERPs were formed from omissions of certainly unpredictable tones and potentially predictable tones from all three experiments. In line with previous experiments, ERPs were baseline corrected using the pre-stimulus time window as baseline window (-150 ms to 0 ms relative to stimulus-onset). In Experiment 1, the repetition reliability was manipulated in five conditions and in Experiment 2 and 4, there was a manipulation of repetition accuracy instead. ERPs were averaged so that the levels 0 %, 25 %, 50 %, 75 % and 100 % of Repetition Reliability in Experiment 1 corresponded with the levels 40 %, 55 %, 70 %, 85 % and 100 % of Repetition Accuracy in Experiment 2 and 4. One might question the validity of this approach because physically different degrees of predictability were taken together to form an average (e.g. 0 % Repetition Reliability [Exp. 1] averaged with 40 % Repetition Accuracy [Exp. 2 and 4]). However, this approach can be justified because the choice of levels of Repetition Accuracy was based on a pilot experiment in which subjects were asked to rate whether they perceived the stimuli as pairs or rather as single tones with random frequency. Because tones were perceived as single events (rather than pairs) to the same extent from 0 % to 40 % accuracy, 40 % was chosen as the lowest value. Hence, the different gradations of Repetition Reliability and Repetition Accuracy might not match physically but perceptually (see Section 2.2.1 for details). For reasons of simplicity, the experimental manipulations of Repetition Reliability and Repetition Accuracy will henceforth be called Predictive Certainty (5 levels: 0 %, 25 %, 50 %, 75 % and 100 %). An overview of the number of stimuli, used to form the ERPs in the respective Experiments, is given in Table 4.1. Statistical analyses were carried out on ERPs obtained from electrode position Cz in Experiment 1 and Experiment 4 and correspondingly from electrode position E01 in Experiment 2. Since electrode position Cz and E01 are located at identical positions on the scalp, in the following, this position will consistently be referred to as electrode position Cz. To investigate ERP responses to omissions of certainly unpredictable vs. potentially predictable tones across different degrees of Predictive Certainty, the statistical analyses were carried out in line with Experiment 1 and 2. A within-subject RMANOVA was conducted with the factor Stimulus Type (2 levels: change certain, repetition possible) and the factor Predictive Certainty (5 levels: 0 %, 25 %, 50 %, 75 % and 100 %) for omissions measured at Cz in the interval of 0 - 50 ms relative to stimulus-onset.

To provide a better overview of the ERP effects across conditions, scalp potential maps of the ERPs in the extreme conditions of each experiment were plotted together in the respective analysis interval. For all RMANOVAs, the Greenhouse-

Geisser correction was applied and the epsilon correction factor is reported whenever the sphericity assumption was violated as indicated by Mauchly's test.

**Analysis of peri-stimulus ERPs of tones** In order to analyze ERP responses to tones across experiments, ERPs of subjects in all Experiments (1, 2, 3 and 4) were pooled together to form grand-average ERPs. This resulted in a sample size of 89 subjects. EEG preprocessing and data analysis were carried out in line with the analysis of the respective Experiment (for details see Section 2.1.1 for Experiment 1, Section 2.2.1 for Experiment 2, Section 3.2.1 for Experiment 3 and Section 3.3.1 for Experiment 4). The resulting ERPs were formed from certainly unpredictable

Table 4.1: Numbers of stimuli per condition for all experiments. Displayed are the numbers of stimuli which were valid to use for further analysis and after systematic exclusion of corrupted epochs (e.g. tones within 600 ms after onset of an omission)

| | condition | Exp. 1 | Exp. 2 | Exp. 3 | Exp. 4 |
|---|---|---|---|---|---|
| $n(tones)$ | 1 | 2996 | 1800 | 1250 | 2880 |
| certainly | 2 | 637 | 1800 | 1250 | 2880 |
| unpredictable | 3 | 1433 | 1800 | 1250 | 2880 |
| | 4 | 2221 | 1800 | 1250 | 2880 |
| | 5 | 3796 | 1800 | 1250 | 2880 |
| $n(tones)$ | 1 | - | 1800 | 1250 | 2880 |
| potentially | 2 | 5355 | 1800 | 1250 | 2880 |
| predictable | 3 | 4556 | 1800 | 1250 | 2880 |
| | 4 | 3771 | 1800 | 1250 | 2880 |
| | 5 | 3796 | 1800 | 1250 | 2880 |
| $n(omissions)$ | 1 | 200 | 240 | - | 144 |
| certainly | 2 | 200 | 240 | - | 144 |
| unpredictable | 3 | 200 | 240 | - | 144 |
| | 4 | 200 | 240 | - | 144 |
| | 5 | 200 | 240 | - | 144 |
| $n(omissions)$ | 1 | - | 240 | - | 144 |
| potentially | 2 | 200 | 240 | - | 144 |
| predictable | 3 | 200 | 240 | - | 144 |
| | 4 | 200 | 240 | - | 144 |
| | 5 | 200 | 240 | - | 144 |

tones and potentially predictable tones from all four experiments. For better comparability, only tones from the 150 ms SOA condition were used from Experiment 3. Since the presence or absence of omissions had no influence on the processing of tones in Experiment 4 (see Section 3.3.2), ERPs of both omission conditions were used for further analysis. An overview of the number of stimuli, used to form the ERPs in the respective Experiments, is given in Table 4.1. Statistical analyses were carried out on ERPs obtained from electrode position Cz. To investigate the influence of Predictive Certainty on the processing of tones and to probe for possible pre-stimulus ERP correlates of prediction, the same analysis strategy was used as described in Section 3.1.1. No baseline correction was performed to avoid carrying over any effects from pre-stimulus to post-stimulus time windows or vice versa, thus allowing for a neutral assessment of tone processing both before and after the onset of a stimulus.

Statistical testing was performed both, for certainly unpredictable and potentially predictable tones. However, it should be noted that only the potentially predictable tones are informative regarding electrophysiological correlates of prediction formation in the brain. The rationale of the statistical analysis was based on finding graded ERP effects of predictability across the five different conditions. This was done by means of linear trend tests as part of within-subject RMANOVAs. Hence in order to test whether predictive certainty influenced tone processing across all experiments, an RMANOVA with the factor Predictive Certainty (5 levels: 0 %, 25 %, 50 %, 75 % and 100 %) was conducted for both the pre-stimulus and the post-stimulus window. In case of significant main effects in the RMANOVA, within-subject linear contrast analyses were performed to test for linear monotonic trends in the data. In line with previous analyses, the processing dynamics throughout the entire epoch were explored. To this aim, a running linear trend test was performed separately for each sampling point and corrected for multiple comparisons using false discovery rate (Benjamini & Hochberg, 1995). Significant intervals are indicated by colored bars at the abscissa of each ERP plot (cf. Figure 4.3). The bars are colored in red if the amplitude values at the respective point were positively correlated with the model coefficients from the linear trend test, and in blue if the ERP values were negatively correlated with the model coefficients. Scalp potential maps of the ERPs in the extreme conditions of each experiment were plotted together. For all RMANOVAs, the Greenhouse-Geisser correction was applied and the epsilon correction factor is reported whenever the sphericity assumption was violated as indicated by Mauchly's test.

## 4.2 Results

**Omission ERPs** Grand-average ERPs of the omissions are shown in figure 4.1. Scalp topographies of the omission ERP modulation by predictive certainty for all experiments containing omissions are shown in Figure 4.2. The 2 x 5 RMANOVA with the factors Stimulus Type and Predictive Certainty for tone omissions yielded a significant interaction of Stimulus Type by Predictive Certainty [$F(4,236)=3.693$, $p=0.00616$, $\eta^2=0.059$] and a main effect of Stimulus Type [$F(1,59)=130.340$, $p<0.00001$, $\eta^2=0.688$]. There was no main effect of Predictive Certainty [$F(4,236)=3.693$, $p=0.00616$, $\eta^2=0.059$]. To resolve the interaction, two within-subject RMANOVAs with the factor Predictive Certainty were conducted for each, certainly unpredictable and potentially predictable tone as follow-up analyses. For certainly unpredictable tones, there was no main effect of Predictive Certainty [$F(4,236)=1.182$, $p=0.31956$, $\eta^2=0.020$] but for potentially predictable tones a significant main effect of Predictive Certainty was observed [$F(4,236)=3.030$, $p=0.01835$, $\eta^2=0.049$] which followed a linear trend [$F(1,59)=8.769$, $p=0.00441$, $\eta^2=0.129$]. This pattern of results indicates that the degree of predictive certainty only modulated the processing of omissions of potentially predictable tones but not of certainly unpredictable tones. ERPs in response to potentially predictable tones showed more positive deflections in association with higher degrees of predictive certainty.

**Peri-stimulus tone ERPs** The point-wise linear trend test indicated a positively correlated linear trend prior to the onset of potentially predictable tones which lasted approximately until the end of the tone (i.e., 50 ms; cf. Figure 4.3). A negative correlation was observed at the beginning and at the end of the epoch of potentially predictable tones (from -150 to -120 ms as well as from 90 to 150 ms relative to tone onset). No ERP modulation by Predictive Certainty was observed from 120 to 25 ms before onset of the potentially predictable tone. The RMANOVA with the factor Predictive Certainty for potentially predictable tones yielded a significant main effect in the pre-stimulus window [$F(4,352)=7.939$, $p=0.00001$, $\eta^2=0.083$, $\epsilon=0.908$] which followed a linear trend [$F(1,88)=21.118$, $p=0.00001$, $\eta^2=0.194$]. This effect indicates that the ERPs shortly before the onset of potentially predictable tones varied depending on the predictive certainty embedded in the tone sequences. Higher forms of predictive certainty are associated with more positive ERP ampli-

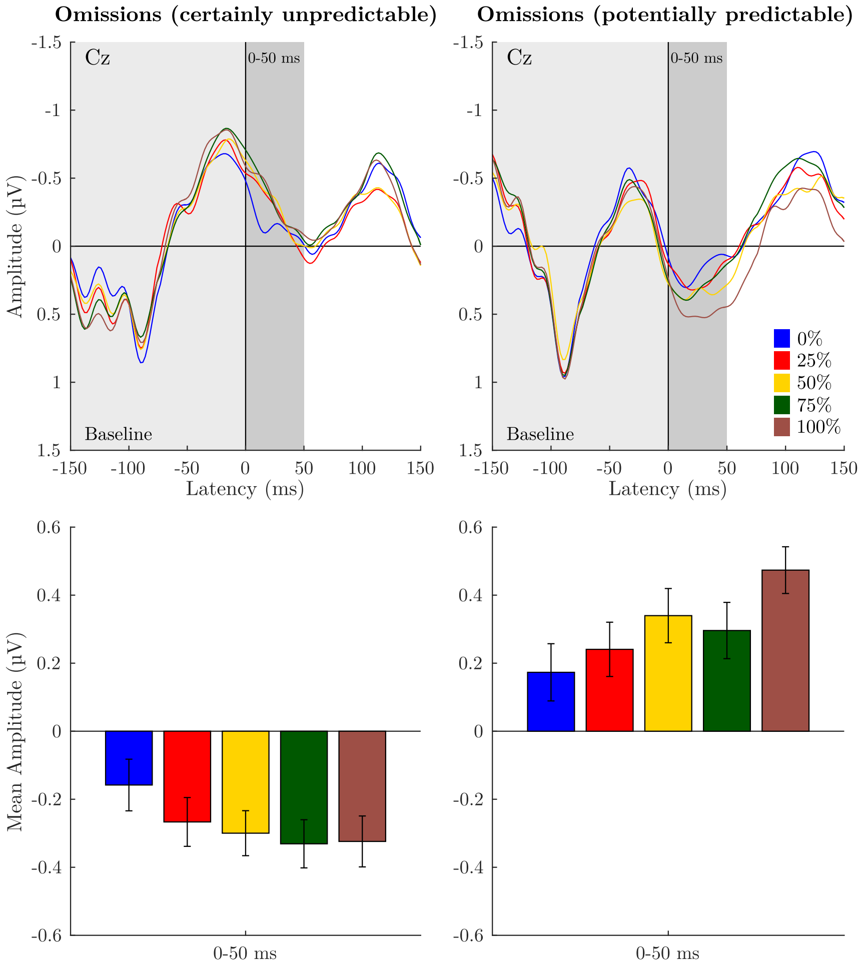


Figure 4.1: Electrophysiological Results. Upper panel: grand-average ERPs of omissions (certainly unpredictable vs. potentially predictable) across all levels of Predictive Certainty averaged across Experiments 1, 2 and 4. Lower panel: bar diagrams with mean amplitude values for each condition in the interval of 0 ms to 50 ms relative to stimulus-onset. Error bars indicate standard errors of the mean.

tudes. This effect remains stable also after the onset of the tone as indicated by a significant main effect of Predictive Certainty for the post-stimulus window [$F(4,352)=19.144$, $p<0.00001$, $\eta^2=0.179$, $\epsilon=0.872$] which also followed a linear trend [$F(1,88)=56.656$, $p<0.00001$, $\eta^2=0.392$]. These effects were also present for certainly unpredictable tones [pre-stimulus: $F(4,352)=13.675$, $p<0.00001$, $\eta^2=0.134$; post-stimulus: $F(4,352)=21.660$, $p<0.00001$, $\eta^2=0.198$, $\epsilon=0.798$] which also followed a linear trend [pre-stimulus: $F(1,88)=16.427$, $p=0.00011$, $\eta^2=0.157$; post-stimulus: $F(1,88)=19.664$, $p<0.00001$, $\eta^2=0.183$]. For certainly unpredictable tones this linear trend was reversed indicating that higher degrees of predictive certainty are associated with more negative ERP amplitudes both, before and after stimulus-onset. Note, that the negative linear trend prior to certainly un-

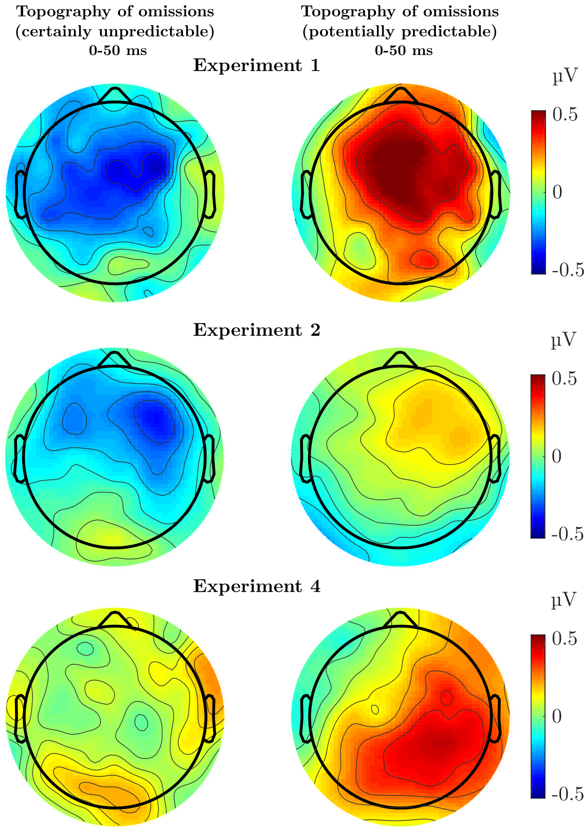


Figure 4.2: Topographical scalp voltage distribution of omission ERP modulation by Repetition Reliability (Experiment 1) and Repetition Accuracy (Experiment 2 and 4) in the interval of 0 ms to 50 ms relative to stimulus-onset. Plotted are the difference waves of the extreme conditions (100 % minus 0 % [Experiment 1] and 100 % minus 4 % [Experiment 2 and 4]).

predictable tones started much earlier then the positive trend prior to potentially predictable tones (at about -60 ms relative to stimulus-onset) and did not last as long as for the potentially predictable tones (at about 40 ms relative to stimulus-onset). Scalp topographies of the ERP modulation by predictive certainty for all experiments are shown in Figure 4.4.

## 4.3 Discussion

The current analysis aimed at reassessing results from the conducted experiments. Data across several experiments was accumulated in order to increase the statistical power. For a reinvestigation of the ERP modulations in response to tone omissions, data from 3 Experiments was merged together resulting in a sample

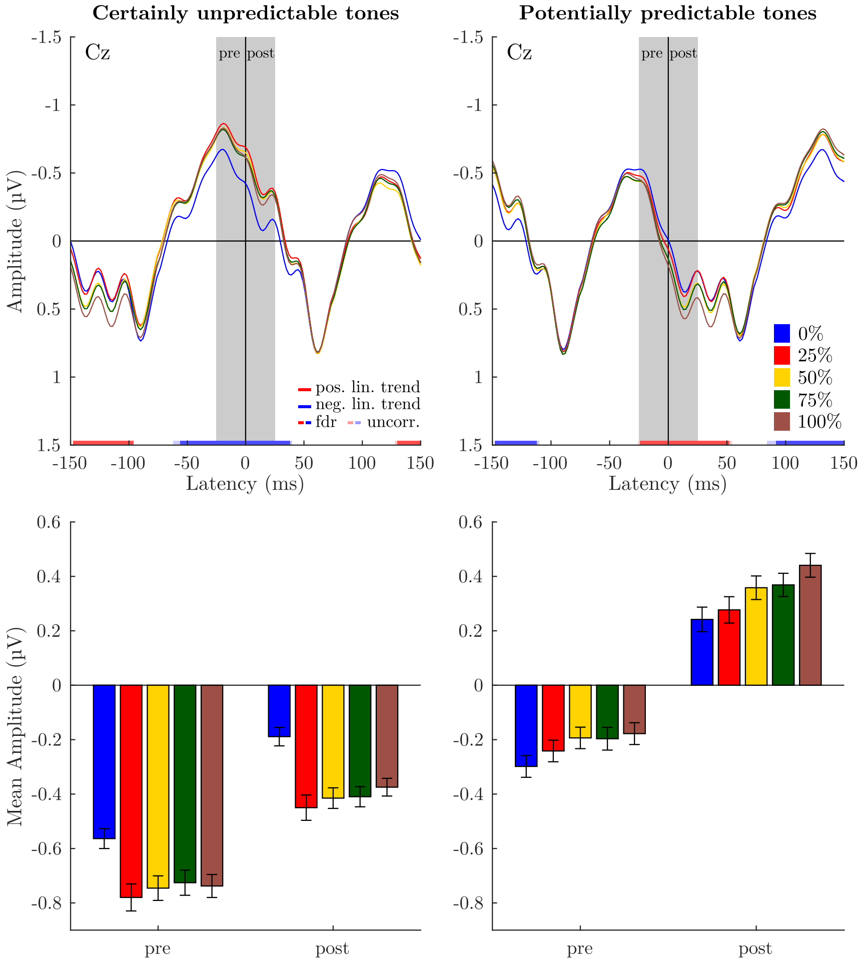


Figure 4.3: Electrophysiological Results. Upper panel: grand-average ERPs of tones (certainly unpredictable vs. potentially predictable) across all levels of Predictive Certainty averaged across all experiments. Lower panel: bar diagrams with mean amplitude values for each condition in the interval of -25 ms to 0 ms (pre-stimulus window) and 0 ms to 25 ms (post-stimulus window) relative to stimulus-onset. Error bars indicate standard errors of the mean.

size of 60 subjects. For the reassessment of peri-stimulus ERP modulations, data of all four experiments was accumulated resulting in a sample size of 89 subjects. The data accumulation was conducted in order to increase the signal-to-noise ratio of the obtained averaged ERP results and to serve as a countermeasure against detecting false positive results and inflated effects sizes. If the initial effects of Repetition Reliability for ERPs in response to potentially predictable tones observed in Experiment 1 are indeed true effects, they should also show up in the grand average ERPs across the included experiments. Likewise, graded effects associated with different degrees of predictive certainty should become apparent at approximately 25 ms prior to the onset of potentially predictable tones if the

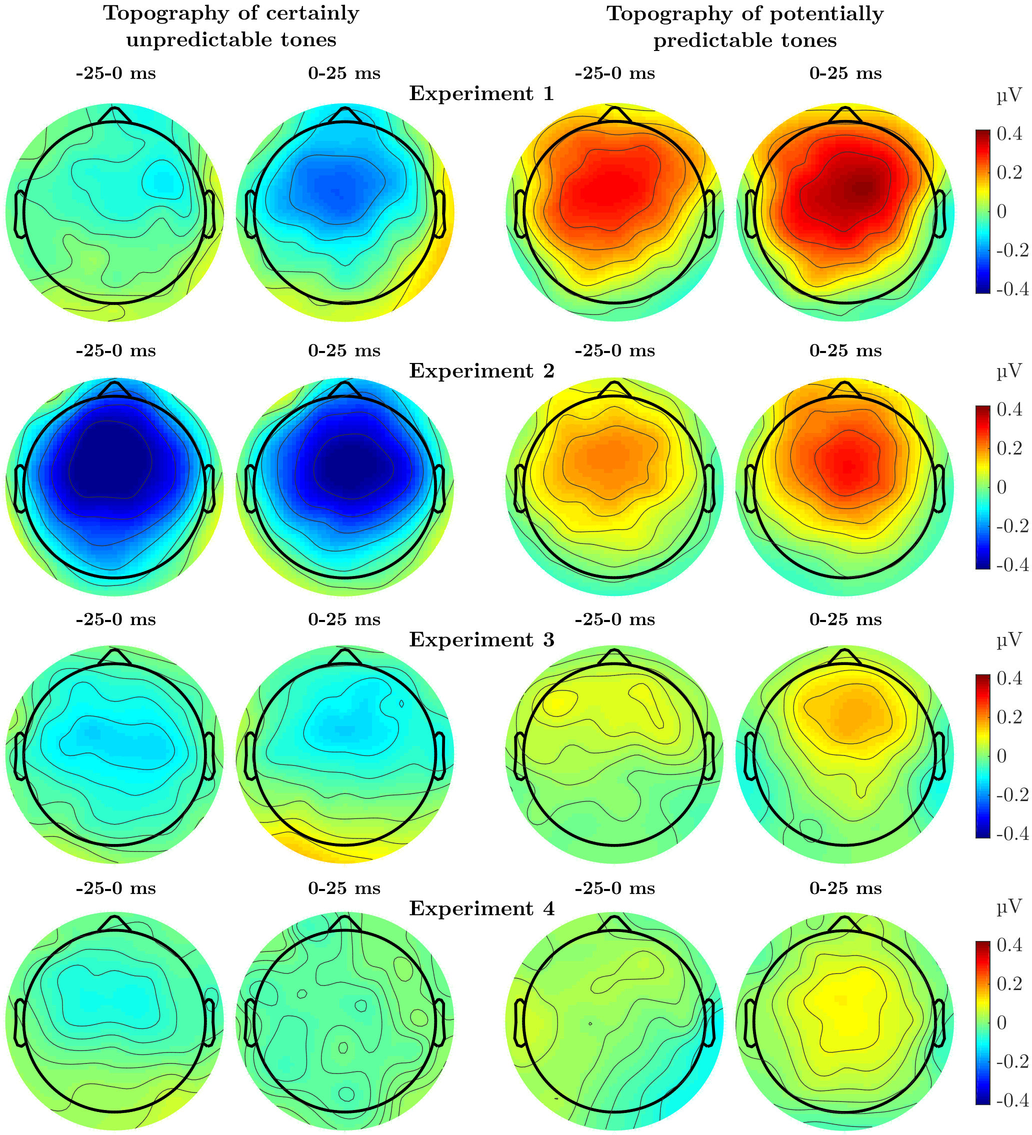


Figure 4.4: Topographical scalp voltage distribution of ERP modulation by Repetition Reliability (Experiment 1) and Repetition Accuracy (Experiment 2, 3 and 4) in the interval of -25 ms to 0 ms and 0 ms to 25 ms relative to stimulus-onset for all 4 experiments. Plotted are the difference waves of the extreme conditions (100 % minus 0 % [Experiment 1] and 100 % minus 40 % [Experiment 2, 3, 4]).

effects observed in Experiment 1 and 2 are not only incidental findings. Moreover, if the unexpected negative linear trend around the onset of certainly unpredictable tones in Experiment 2 persists in the accumulated average, it might indeed be a meaningful effect even though it was just observed in one of four experiments.

The results for the accumulated ERPs in response to omissions revealed a significant interaction of Stimulus Type by Predictive Certainty. Follow-up tests suggested a main effect of Predictive Certainty for potentially predictable tones but not for certainly unpredictable tones. This suggests that the observed effect for ERPs of potentially predictable tones in Experiment 1 might indeed reflect a true effect and that the SNR in the remaining two experiments might have been insufficient. Likewise, results of the peri-stimulus effects of ERPs, accumulated across all experiments, revealed a positive linear trend in the range between -25 ms to 50 ms relative to the onset potentially predictable tones. Like in the first two experiments, this effect was preceded by a relatively long period in which no ERP modulations by predictive certainty were observed (-110 ms to -25 ms relative to stimulus-onset). These results reveal that the effects observed in Experiment 1 and 2 survived even in the large accumulated dataset with 89 subjects. This provides evidence in favor of graded pre-stimulus effects associated with different degrees of predictive certainty. However, the negative linear trend around the onset of certainly unpredictable tones, observed in Experiment 2, also survived in the accumulated results. Lower degrees of predictive certainty incorporated in the respective condition were associated with more positive deflections in the ERPs in the range between -60 ms to 40 ms relative to stimulus-onset. Note that this is the reversed pattern of results observed around the onset of potentially predictable tones. Note also that the direction of the effect is mainly driven by the condition with the lowest degree of predictive certainty. Like in Experiment 2, this effect starts much earlier than the positive linear trend observed in response to potentially predictable tones. This suggests that the effect might more likely be driven by carry-over effects from previous tones.

In general, the current results suggest that the effects observed in the single experiments cannot simply be dismissed as incidental findings. The remaining question is whether the effects showed up only because they were driven by the existing effects from Experiment 1 and 2 or whether data from the other two experiments also contributed to this effect. It might be that true effects in Experiment 3 and 4 were not detected in isolation due to insufficient data quality, but with higher statistical power the effects might have further consolidated the effects that were already present in the first two experiments. Alternatively, there was no ERP modulation by predictive certainty in Experiment 3 and 4 and the effects from Experiment 1 and 2 were strong enough to survive in the accumulated dataset. With the current set of results, it is virtually impossible to answer this question. Even if true effects were present in all of the conducted experiments, it is still difficult to draw a clear conclusion about the true nature of the results. Graded effects of ERP modulations for both, potentially predictable and certainly unpredictable tones were observed. However, the certainly unpredictable tones were considered as a control condition for which no signs of prediction-related activity were expected. If a parametric divergence of ERPs is really associated with the predictive certainty of an event, there should not be such an effect in response to certainly unpredictable tones.

Moreover, the current pattern of results does not provide enough information to differentiate whether the observed ERP modulations indeed reflect prediction-related effects or whether they are driven by carry-over effects from previous tones. Experiment 3 was designed to resolve this issue but as we have seen, no significant effects of Repetition Accuracy on ERP amplitudes (neither for pre-stimulus nor for post-stimulus ERPs) were observed in this dataset. In order to unequivocally resolve this issue, further research is urgently needed. In the following part, the results of the current thesis will be discussed in a broader context and recommendations for future investigations will be provided.

# 5 General Discussion

The current thesis aimed at systematically investigating predictive auditory processing and the influence of different degrees of predictive certainty on prediction-related electrophysiological measures using EEG. Based on Bendixen et al. (2009), four experiments were performed following the same basic logic: subjects were presented with an isochronous sequence of tones in which predictive relations between successive tones varied systematically across five different levels. In the first experiment, predictive certainty was manipulated by varying the repetition probability in five conditions. In the remaining three experiments, the repetition accuracy was manipulated. Results from one experiment demonstrate graded effects of predictive certainty for ERPs in response to omissions of potentially predictable tones. However, this pattern of results could not be replicated in two further experiments. Furthermore, graded effects of predictive certainty were observed in two experiments, starting shortly before the onset and lasting up until the end of potentially predictable tones. This indicates that the brain shows signs of predictive processing mediated by different degrees of predictive certainty even before the start of a potentially predictable event. Note that this effect could not be replicated in two follow-up experiments which were designed to further characterize the proposed underlying mechanism. However, graded effects of predictive certainty could be regained after collapsing the data across experiments, both for omissions of potentially predictable tones and also shortly before until the end of potentially predictable tones. This result strengthens the notion that the graded prediction-related pre-stimulus modulations indeed reflect true effects of predictive auditory processing. However, no clear conclusions can be drawn due to some remaining inconsistencies in the results, like the reversed effect in the pre-stimulus ERPs of certainly unpredictable tones. As introduced in Part 2, a vast number of studies provides indirect evidence in favor of predictive processing by demonstrating effects associated with prediction error processes. However, results demonstrating direct evidence of predictive auditory processing, like prediction-related activity that unfolds prior to the onset of a predictable event is currently lacking. In combination with the inconsistencies of the current results, this raises questions about the suitability of measures like ERPs to validly investigate "purely" prediction-related correlates of predictive auditory processing.

Recently, Sedley et al. (2016) recorded local field potentials (LFPs) on patients suffering from epilepsy using Electrocorticography (ECoG) during neurosurgery. They recorded oscillatory activity at the surface of the neocortex in response to sequences of tones, partitioned into segments of variable length. Each segment was randomly drawn from a Gaussian distribution, characterized with a mean frequency and a standard deviation. Therefore within one segment, stimuli were partly predictable and upon a change of segments, prediction violations were induced and the predictive model had to be adjusted. As in the current series of experiments, this approach was designed to investigate neural responses to manipulations of predictability of stimulus repetitions under uncertainty. The authors furthermore aimed at finding direct evidence for different aspects of predictive processing, like activity that is linked to prediction error, surprise, the updating of predictions and the precision of predictions. Sedley et al. (2016) observed associations between different parameters of predictive coding with different frequency bands of oscillatory brain responses. Surprise due to prediction violations was linked to oscillations in the gamma-band (>30 Hz) which is, among others, associated with the unexpectedness of incongruence of stimuli (Arnal, Wyart, & Giraud, 2011; Brodski, Paasch, Helbling, & Wibral, 2015; Todorovic et al., 2011). Changes in predictions were associated with oscillations in the beta-band (12-30 Hz) which are usually associated with motor actions (Neuper & Pfurtscheller, 2001), cognitive processes like memory rehearsal (Tallon-Baudry, Bertrand, & Fischer, 2001) and which are also indirectly linked to prediction (Arnal & Giraud, 2012). From the perspective of the current thesis, the most interesting results were the associations between prediction precision and oscillatory activity in the alpha-band (8-12 Hz). The alpha-band has been associated with a variety of cognitive processes, like memory (Klimesch, Schimke, & Pfurtscheller, 1993; Klimesch, 1997) and attention (Yordanova, Kolev, & Polich, 2001). Alpha-magnitude has been found to correlate with the probability of a stimulus occurring (Bauer, Stenner, Friston, & Dolan, 2014) and has been demonstrated to modulate higher frequency oscillations through phase-amplitude coupling (Jensen & Mazaheri, 2010). Critically for the current thesis, the oscillatory correlates of prediction precision were mainly induced alpha-activity. As opposed to evoked oscillatory activity, induced oscillatory activity is only correlated with an experimental manipulation but not strictly time-locked to the onset of a stimulus (Herrmann, Grigutsch, & Busch, 2005).

However, only evoked oscillatory activity might be observed in ERPs, whereas induced oscillatory activity is assumed to be cancelled out by the averaging technique that is applied when computing event-related potentials. In the worst case, this could imply that the current approach of utilizing ERPs is not suitable for investigating predictive processing under different degrees of predictive certainty.

## 5.1 Implications for current research

The results observed by Sedley et al. (2016) imply that certain electrophysiological correlates of auditory prediction, related to different degrees of predictive certainty, might be difficult, if not even impossible, to isolate using event-related potentials. This would have major implications for the current results and previous publications investigating predictive auditory processing. It would furthermore explain why the expected effects in the current thesis are so small or even absent in replication experiments despite using identical experimental manipulations.

In the introduction of the current thesis, several different studies investigating predictive processing in audition were introduced. Different experimental approaches were categorized, like match-paradigms (see, e.g. Baldeweg, Klugman, Gruzelier, & Hirsch, 2004; Costa-Faidella, Grimm, et al., 2011; Haenschel et al., 2005), mismatch-paradigms (see, e.g. Bendixen & Schröger, 2008; Paavilainen et al., 2007; Grimm et al., 2011; Widmann et al., 2007), and omission-paradigms (see, e.g. Bendixen et al., 2009). All of the established paradigms build up some form of predictability in order to investigate electrophysiological correlates of different aspects of predictive processing, like neural markers of top-down prediction or prediction error signals. Furthermore, all of these approaches share some common methodological problems which make it difficult to unequivocally interpret the obtained results in terms of predictive coding. For example, all of the presented studies use post-stimulus potentials in order to investigate a process that should theoretically be observable already before the onset of potentially predictable events. Therefore, it cannot be ruled out that the observed results might reflect the outcome of a retrospective mechanism instead of being correlates of a proposed predictive process. Apart from mismatch-paradigms which by definition have to make use of post-stimulus potentials, experimental investigations of predictive processing would hugely benefit from extending their scope by also examining what happens in the pre-stimulus time range. Another major drawback of many of the presented studies is that they only compare fully predictable with completely unpredictable conditions. This has some serious implications regarding the ecological validity of the results. Natural signals incorporate different degrees of predictive certainty, and more often than not they fall short of perfect predictability. The brain is sensitive to changes in predictability of sensory signals (Nastase, Iacovella, & Hasson, 2014; Tobia, Iacovella, & Hasson, 2012; Tobia, Iacovella, Davis, & Hasson, 2012; Tremblay, Baroni, & Hasson, 2013). Highly predictable signals would tend to be linked to technical noise (e.g. the buzzing of a refrigerator or the air conditioning), while natural signals (like the babbling of a brook or the sound of the sea) are much more variable - but still predictable to some extent. It is not evident how findings from classical studies, contrasting fully predictable with fully unpredictable stimuli, apply to processing of such more natural signals. Furthermore, in order to decrease the likelihood of incidental findings, it is important to probe more than two conditions, creating the opportunity to find links between a gradual variation of predictive relations and correspondingly graded ERP effects.

One major aim of the current thesis was to find indicators of "true" predictions being generated in the brain - that is, to see modulations of brain responses by predictability before rather than after the onset of sensory events. Electrophysiological correlates of auditory predictions systematically varied with the degree of predictive certainty embedded in the stimulus streams. In other words, the system shows signs of predictive processing even when the predictive information is not fully reliable (i.e., a regularity is not always applicable) and also when the predictive information turns out to be not fully accurate. The fact that the magnitude of the electrophysiological responses was linearly correlated with the degree of predictive certainty is strong evidence that the underlying predictive mechanism flexibly adapts to the degree of predictive certainty in the sensory input.

Such effects have been shown for other correlates of auditory processing, like the MMN. Sussman and Winkler (2001) presented subjects with short standard tones in an oddball paradigm with occasional frequency deviants. They presented either single deviants or double deviants (two deviants of equal frequency presented successively). They showed that the MMN elicited by the double deviants changed depending on the context in which those deviants were presented. Double deviants elicited multiple MMNs when presented around single deviants but only one MMN when presented in a context of mainly double deviants. Importantly, these context adaptations of the MMN took place within a time range of less than a minute. The authors interpret their findings as a dynamic process of sensory updating which is continuously running in the auditory system. Approaches like the one applied by Sussman and Winkler (2001) convincingly demonstrate that the predictability embedded in the sensory context systematically influences the obtained ERPs. However, due to the nature of post-stimulus ERPs, these correlates do not provide means to tap into 'purely' prediction-related activity or to measure prediction in a literal sense.

Fortunately, this does not affect the pre-stimulus effects because there was no physical stimulation shortly before the onset of the tones. Therefore, the current study provides arguments that pre-stimulus effects permit safer conclusions in terms of predictive processing. The fact that the post-stimulus ERPs show the same polarity and similar topographical distributions as the physically uncontaminated pre-stimulus ERPs suggests that ERP effects after the onset of a stimulus are also indicative of predictive processing in the current study. The time course of the observed effects excludes a number of alternative explanations. The relatively long non-significant interval prior to the observed pre-stimulus response argues against the notion that the observed effects can be explained by simple carry-over effects from prior stimulation. Note however, that carry-over effects cannot clearly be ruled out using this logic because processing differences of prior stimuli might unfold at later stages which, just by coincidence, might be around the onset of the next tone. Experiment 3 was designed as a control for this ambiguity by systematically manipulating the SOA across different conditions. However, since the effects could not be replicated in Experiment 3, it cannot be precluded that the observed pre-stimulus effects might indeed be attributed carry-over effects. Nonetheless, the linear trend mediated by the degree of certainty embedded in the experimental condition was present only around the predictable events and was reversed shortly before and after the onset of certainly unpredictable tones (the tones in between the potentially predictable ones). Within this positively correlated interval, linear trends with identical directions and very similar topographical distributions were observed before and after stimulus-onset. Furthermore, this pattern of results occurred across two experiments with different ways of manipulating predictive certainty. However, a similar effect was observed around the onset of certainly unpredictable tones in Experiment 2 which was reversed in polarity. If graded effects associated with different degrees of predictive certainty are indeed valid indicators of underlying neural processes which are based on predictive coding, these effects should theoretically be absent in response to (and likewise in anticipation of) fully unpredictable events. Despite a much earlier onset of this effect, which at least partially suggests a different origin of the observed ERP modulations, it raises questions about the validity of the observed results.

A related problem was already addressed by Bendixen, Duwe, and Reiche (2015). The authors report an experiment which was based on the tone-pair paradigm by Bendixen et al. (2009) but was extended to a noise-based version of the original paradigm in order to investigate the ecological validity of predictive auditory processing. Specifically, the study aimed at investigating ERPs in response to noise occlusions of either predictable or unpredictable tones. They applied a subtraction approach in order to examine the similarity between the tone-related ERPs and the ERPs elicited "behind" the noise. The authors found a higher degree of similarity in the processing of predictable tones and their unexpected occlusion by noise than in the case of unpredictable tones. However, they also found that differences between ERPs in response to noise occlusions of unpredictable vs. predictable tones was mainly driven by differences between the tone ERPs rather than between the behind-noise ERPs. Bendixen et al. (2015) conclude that these results highlight an important methodological aspect of using the tone-pair paradigm in order to investigate predictive auditory processing. In the tone-pair paradigm by Bendixen et al. (2009), the predictable tones (2nd position) were frequency repetitions while the unpredictable tones (1st position) were frequency changes. Systematic effects observed between these two categories of stimuli might reflect repetition suppression (Baldeweg, 2006; Boutros, Gjini, Eickhoff, Urbach, & Pflieger, 2013; Todorovic & de Lange, 2012). The authors suggest another alternative explanation for the observed pattern of results which is based on grouping effects. Certainly unpredictable tones (1st position) constitute the beginning of a tone pair, whereas potentially predictable tones (2nd position) constitute the end of a tone pair. These tone pairs can be considered auditory objects themselves which might introduce additional processing stages on top of the expected prediction-related processes. The tone-pair paradigm in its original form or in the form applied within the context of the current thesis lacks the ability to clearly disentangle contributions to ERP modulations from different processes like explained above. Therefore, appropriate control conditions need to be applied in order to draw clear conclusions from the observed effects.

It can be concluded that investigating pre-stimulus ERPs in order to assess predictive coding in audition poses several advantages over more traditional approaches since they investigate possible underlying processes more directly and are less prone to be influenced by other coinciding phenomena, like different states of refractoriness. However, several practical and methodological improvements can be made both, for established indicators of predictive auditory processing and for newer alternative indicators. In the following section, some general recommendations will be provided to enhance future investigations of predictive auditory processing and to maximize the amount of information gained by the applied procedures.

## 5.2 Recommendations for future investigations

In the last section, electrophysiological markers of predictive auditory processing have been discussed. Established prediction-related correlates have been compared to alternative indicators of predictive coding in audition, like the pre-stimulus potentials investigated in the context of the current thesis. Several practical and methodological problems have been identified both for more established, as well as for newer alternative electrophysiological markers of predictive auditory processing. Based on these insights, the current section attempts to provide some recommendations for future investigations of predictive coding in audition.

### 5.2.1 Recommendations for the design of experiments

Making decisions about the concept of a planned experiment and designing a detailed implementation which helps to appropriately answer the research question is arguably one of the most important steps in the scientific process. When designing experiments which are aimed at investigating predictive auditory processing, several aspects have to be taken into account.

Aside from the inconsistencies in the observed results, the current thesis improves the original paradigm by Bendixen et al. (2009) in several methodologically important ways. Different aspects of predictive certainty (predictive reliability and predictive accuracy) have been introduced which enables to systematically investigate the ecological validity of proposed theories of predictive coding in audition (i.e., whether it is still functional even if the stimuli cannot be predicted with full certainty and accuracy). As laid out in the introduction, previous studies investigating predictive auditory processing often implement straight-forward stimulation protocols like oddball paradigms (for a review, see: Schröger, 2007). However, there are also studies investigating predictive auditory processing under more realistic conditions (as compared to rather simple oddball paradigms) with the aim to, first, rule out certain confounding factors associated with simpler stimulation protocols (e.g. refractoriness due to repetitions of physically identical stimuli) and second, to provide further insights into the ecological validity of the proposed underlying mechanism. For example, this has been achieved by using roving paradigms in which the physical features of the standard and deviant stimuli perpetually change within one experimental block (e.g. Baldeweg et al., 2004; Haenschel et al., 2005) or by using more complex stimulation protocols employing abstract rules, conveyed by a conjunction between two stimulus features, like e.g. tone pitch and tone intensity (Paavilainen, Simola, Jaramillo, Näätänen, & Winkler, 2001). However, when using such protocols, rules are still either certainly confirmed or clearly violated. In other words, apart from the complexity of the rules, there is a clear distinction between rule confirmation and rule violation which is rarely the case in natural signals. Here, the current thesis offers some clear improvements by providing means of investigating predictive processing under the influence of different degrees of predictive un-/certainty. This enables to gain further insights into how the proposed underlying mechanism might work under realistic conditions.

Arguably one of the most significant advantages of the applied paradigm over previous approaches, which only contrast fully predictable with certainly unpredictable stimuli, is the gradual manipulation of predictive certainty across five different levels. This step-wise manipulation provides means to gain further insights into the characteristics of the proposed predictive mechanism, like the minimum level of predictive certainty that is necessary for the system to be triggered, or whether higher degrees of predictive certainty are associated with a decrease or an increase in the observed measures. This is of particular importance when no explicit hypotheses exists regarding the direction of the expected effects. Investigations addressing predictive auditory processing might observe systematic effects between predictable and unpredictable stimuli but only limited conclusions can be drawn about the underlying mechanism by using such extreme contrasts. For example, Barascud, Pearce, Griffiths, Friston, and Chait (2016) presented continuous sequences of sine tones. Tones were arranged either regularly or randomly regarding their pitch, and within one block the sequence could change from random to regular and vice versa. The authors measured the subjects' MEG while being passively exposed to the auditory stimulation. Additionally the subjects were engaged in a visual task. Upon a transition from random to regular sequences, the authors observed a sustained response in root mean squared (RMS) MEG power starting at about 100 ms after the transition. The RMS MEG response to the transition was characterized by a gradual increase in amplitude, followed by a subsequent plateau. For the transition from random to regular, the authors observed an initial increase of RMS power, followed by a steep decrease and a subsequent plateau. The authors interpret this sustained-response amplitude as the brain's response to stimulus predictability. As noted by the authors, in previous investigations increasing stimulus predictability is often associated with a decrease in sensory response (Garrido, Sahani, & Dolan, 2013; Wacongne et al., 2011; Bendixen et al., 2009; Garrido, Kilner, Kiebel, & Friston, 2009) which is often interpreted as suppression of prediction error. However, the authors observed an increased response associated with higher stimulus predictability. Barascud et al. (2016) draw the conclusion that the observed responses reflect precision-weighted sensory signals. They argue that sensory signals associated with low uncertainty (high precision) are indicative of salient sensory evidence which in turn leads to heightened sensitivity (increased gain). Following this logic, a decreased response associated with higher predictability can be interpreted as suppression of prediction error and an increased response can be explained as increased gain. This example shows that any observed change between two conditions, irrespective of the direction, can be interpreted within the context of predictive coding. The only possible outcome arguing against predictive coding would be no effect at all between such two conditions. Hence, using only two extreme conditions might not be sufficient to clearly dissociate between effects of predictive processing and alternative variations in the data. The gradual manipulation of predictive certainty, applied in the current thesis, successfully acts as a countermeasure against such caveats. More levels of manipulation increase the number of possible patterns that can be observed and decreases the risk of finding the exact pattern that was expected just by coincidence. For example, the effects of predictive reliability (Experiment 1) and predictive accuracy (Experiment 2) on ERPs prior to the onset of potentially predictable tones, observed in the current thesis, could not be replicated in two follow-up experiments. This raises the question whether the observed effects are merely incidental findings. However, as noted above, the applied paradigm drastically minimizes the risk of producing incidental findings. Out of 120 different possible combinations, a meaningful pattern of results was observed which was highly similar across two independent experiments regarding its polarity, latency and topographical scalp distribution. Furthermore, such a gradual manipulation can enormously increase the amount of information gained by the obtained results. This can be extremely helpful to provide information about the proposed underlying mechanism. For example, does it vary as a function of the degree of manipulation? Does a certain degree of manipulation has to be exceeded in order to trigger the underlying system? Will the effect increase linearly or quadratically or according to an all-or-none principle? Questions like these simply cannot be answered using only two levels of manipulation.

As an additional measure, the expected effect sizes should be taken into account when planning a new study. Due to several reasons which were discussed in detail in Part 4, many published research findings are based on data from relatively small samples (see, e.g. Pannese, Herrmann, & Sussman, 2015; Sussman & Winkler, 2001; Yabe et al., 1997). When aiming at investigating pre-stimulus correlates of auditory prediction, the current results suggest that the observed effects are usually very small and difficult to access with poor data quality or an insufficient SNR. This doesn't necessarily mean that it is required to acquire an unreasonably big sample size in terms of the number of subjects. Depending on the signal-to-noise ratio that can be achieved with the resources at hand, results from the current thesis suggest that sample sizes of about 20 to 30 subjects might already be sufficient. However, it is always recommended to estimate the required sample size before starting data collection based on the expected effect sizes in order to achieve sufficient statistical power (see, e.g. Faul, Erdfelder, Lang, & Buchner, 2007). The more influential factor might be the number of trials presented per condition that are used to form the averaged potentials (Luck, 2005). Moreover, the SNR of an averaged measure is not linearly related to the number of measurements. The expected SNR is assumed to be increased in proportion to the square root of the number of measurements (Vaseghi, 2013). This is important to bear in mind because the number of presented stimuli needs to be increased substantially in order to achieve a desired increase in SNR. Most of the time this isn't feasible within one experimental session especially when using multiple different degrees of manipulation within one experiment. One possible solution might be to split the experiment into two ore more sessions. In EEG measurements, the expected between-subject variability is usually much higher than the within-subject variability between several sessions. Based on these general considerations, a specific suggestion for a suitable follow-up to the experiments of the present thesis is developed in this section.

Experiment 3 was designed to investigate temporal dynamics of predictive auditory processing and to further rule out alternative explanations of the peri-stimulus effects observed in Experiment 1 and 2. Considering the current state of results, a reasonable next step would be to conduct a replication of Experiment 3 since it includes one condition (150 ms SOA) which is basically also a replication of Experiment 2 and would provide further new information about the temporal dynamics of the observed effects which is lacking in the current results. In order to improve the signal-to-noise ratio, the number of presented stimuli should be adjusted at least to the number of stimuli in Experiment 2. To better replicate the experimental conditions, a reasonable consideration would be to include rare stimulus omissions, like in the first two experiments. Both of these modifications would drastically increase the duration of data acquisition which might introduce unreasonable measuring conditions. This problem might be solved by splitting the experiment into several sessions (e.g. one session for each SOA condition). The SOA condition could alternatively be treated as between-subject factor by testing different groups of subjects for each level of SOA. However, due to relatively strong inter-individual differences between EEG measurements, it is recommended to choose the within-subject solution by splitting the data acquisition for one subject into several sessions. If the effects observed in Experiment 1 and 2 could be regained, it would confirm their validity and provide further information about the temporal dynamics of the obtained electrophysiological measures. Moreover, this experiment would provide essential measures which are necessary to resolve some of the major confounding factors that are present in the conducted experiments by systematically disentangling ERP contributions from the current stimulus and from previous stimuli which in turn would help to explain the reversed peri-stimulus effects for certainly unpredictable tones in Experiment 2. Furthermore, if the onset of the effects observed around potentially predictable tones turns out to be unaffected by the SOA, this would favor the predictive account. Likewise, if the onset of the effect changes with the SOA, this might suggest that the observed ERP-modulations are carry-over effects from previous tones.

Note that apart from the abovementioned improvements, there are also more elementary problems that need to be taken into account when designing an experiment to investigate predictive auditory processing. For example, many traditional paradigms, like the oddball-paradigm, rely on frequently presenting standard stimuli and occasionally presenting deviant stimuli. Such a stimulation protocol in and of itself has an effect on certain ERP components, like the N1 and/or the MMN due to differential states of refractoriness of underlying neural populations, sensitive to the features of the standard vs. the deviant stimulus (cf. Schröger, 2007;

Näätänen & Alho, 1997). Schröger and Wolff (1996) designed a control procedure that enables to estimate the amount of refractoriness-driven contributions to ERP-components in response to a specific stimulus. They embedded deviant stimuli in control blocks that consist of stimuli varying with respect to the same dimension that defines standard and deviant stimulus. Furthermore, each stimulus in the control sequence occurs equally often (identical to the probability of the deviant in the oddball task). The magnitude of the response to deviants in the control condition provides an estimate about the neural contributions that are mainly driven by refractoriness. Furthermore, ERP-components that are commonly associated with predictive processing might be modulated by attention (e.g. Hillyard, Hink, Schwent, & Picton, 1973), task-relevance (e.g. Schröger & Wolff, 1998a, 1998b) or by the rate of presentation, like in the case of N1 attenuation (Sussman & Winkler, 2001) or repetition suppression of the P50 component (Baldeweg, 2006; Boutros et al., 2013; Todorovic & de Lange, 2012). Therefore, it is of particular importance to design appropriate control conditions in order to be able to rule out alternative explanations which would further question the validity of conclusions drawn from the results. For example, when using a tone-pair paradigm, like one introduced by Bendixen et al. (2009), a control condition should be implemented as suggested by Bendixen et al. (2015) in order to rule out that effects might be solely explained by additional processes which are not related to predictive processing, like auditory object formation.

### 5.2.2 Recommendations for data analysis

Apart from the theoretical aspects discussed in the last section, several practical issues have to be taken into consideration when planning and conducting data acquisition. Despite the fact that a large variety of software as well as highly sophisticated signal processing algorithms are openly available today which fundamentally help to increase the signal-to-noise ratio of the measured data, nothing can replace relatively clean and artifact free raw data. Many procedures commonly applied during EEG-data preprocessing, like filtering or ICA-decomposition can drastically alter the results by applying transformations based on certain assumptions about the data. However, if these assumptions are not met, the procedures mentioned above can also distort the measured data to the point that certain effects get inaccessible that otherwise would have been observable. Or even worse, that spurious effects come into existence due to the applied signal-processing tools.

For example, the application of frequency filters is a ubiquitous step in EEG data preprocessing. Widmann, Schröger, and Maess (2015) argue in favor of using frequency filters for ERP research because they have the capability of significantly increasing the SNR (for a more restrictive perspective on filtering, see: Luck, 2005). However, the authors also issue a warning because poor filter design might lead to unintended signal distortions, like a systematic underestimation of the onset latency of ERP components (VanRullen, 2011), artificial components (Acunzo, MacKenzie, & van Rossum, 2012) or spurious dependencies of stimulus detectability on pre-stimulus phase (Zoefel & Heil, 2013). Some of these problems are caused by particular types of filters, like a phase-delay of causal filters which can influence the onset latencies of obtained ERP components. Such problems can partly be circumvented by instead using zero phase-shift acausal filters. However, even if such problems are considered, poor filter design can lead to severe signal distortions. For example, high-pass filtering can seriously distort the signal when applying a cutoff frequency above 0.1 Hz (Luck, 2005). Acunzo et al. (2012) systematically showed that late, slow components can induce systematic bias on earlier components of higher frequency that translate into statistically significant ERP modulations when using acausal filters with a cutoff frequency above 0.1 Hz. Such distortions might have also affected studies investigating predictive auditory processing. For example, Schwartze, Farrugia, and Kotz (2013) used an oddball paradigm with either regular or irregular temporal structure to investigate ERP correlates of formal and temporal predictability. They observed increased P50 amplitudes in response to deviants within regular temporal structure and both, increased P50 and N1 amplitudes in response to deviants within irregular temporal structure. The authors argue that their results confirm that the P50 and the N1 amplitudes reliably encode formal and temporal predictability. However, the authors used a band-pass filter with a low-cutoff (high-pass) frequency of 5 Hz (compare to the maximum high-pass cutoff frequency of 0.1 Hz, recommended by Luck, 2005). Due to this high cutoff frequency, the N1 or later components might already have been affected by the filter. As described by Acunzo et al. (2012), this can lead to a systematic bias of preceding, more transitory components. It cannot be ruled out that this might have led to an overestimation of the P50 components observed by Schwartze et al. (2013). Distortions like these might have major implications on a vast number of published results. As a countermeasure, filters should be implemented or audited by trained individuals and it should always be reassessed in what way the applied filters altered the data. For example, this can be done by comparing the filtered data with the unfiltered data. The difference between the filtered data and the raw data reveals the actual portion of the signal that has been affected by the filter and therefore provides an estimate about possible influences of the applied filter on the observed results. For an in-depth reading addressing digital frequency filters, see Widmann et al. (2015).

Another important point to consider is whether commonly used preprocessing techniques can still be applied when investigating alternative correlates of predictive auditory processing. For example, when investigating pre-stimulus potentials, the common technique of using the pre-stimulus ERP interval for baseline correction renders any possible effects within this interval inaccessible. Moreover, applying baseline correction in this case might carry over effects from the pre-stimulus interval into the post-stimulus interval which might furthermore distort the ERP results after stimulus-onset. As discussed in the introduction, many studies seek to identify correlates of predictive auditory processing which unfold very early relative to the onset of an experimental stimulus because such correlates are assumed to be less influenced by later processes (like e.g. attention) and hence, are assumed to better reflect predictive auditory processing (see, e.g. Bendixen et al., 2009; Grimm et al., 2011). However, the results observed in the current thesis partly suggest that such early correlates might already start to unfold prior to the onset of potentially predictable stimuli which vary systematically with the degree of predictive certainty embedded in the sensory environment. As mentioned above, this might have serious implications when applying baseline correction since there might already be systematic variations within the baseline interval which would than interfere with the results observed after stimulus-onset.

Apart form the applied preprocessing techniques, there are several aspects to consider when obtaining electrophysiological data with EEG in order to investigate correlates of predictive auditory processing. The results of the current thesis suggest that manipulations of predictive certainty might only elicit very minute modulations in the obtained ERPs. It is therefore all the more important to keep interfering noise (e.g. line-noise) as low as possible. However, there are countless sources of non-neural interference that are very difficult to eliminate during data acquisition, like low frequency electrodermal activity due to variations of skin conductance driven by stronger activity of perspiratory glands or drying of electrolyte gel due to thermal changes and skin contact (e.g. Hennighausen, Heil, & Rösler, 1993; Vanhatalo, Voipio, & Kaila, 2005; Tallgren, Vanhatalo, Kaila, & Voipio, 2005). Furthermore, the humidity and temperature prevailing in the measurement environment might drastically decrease the SNR of the recorded data, especially with high electrode impedance (Kappenman & Luck, 2010). A very efficient procedure that is able to deal with artifacts like the ones mentioned above, is independent component analysis (Bell & Sejnowski, 1995). When applied correctly, stereotypical artifacts that are elicited by blinks, eye-movements or cardiac activity can be identified and removed from the data. For a detailed description and general guidelines for the application of ICA for artifact-correction, see Ullsperger and Debener (2010).

## 5.3 Future prospects

The results obtained from the experiments, carried out in the scope of the current thesis, were partially inconsistent and effect sizes were in general relatively small. Two experiments which were designed to replicate the observed effects and to further rule out alternative explanations did not succeed. Therefore some important questions remain unanswered in terms of the validity and conclusions that can be drawn from the obtained results. Two experiments provided evidence in favor of predictive auditory processing by demonstrating ERP-modulations associated with different degrees of predictive certainty shortly before, up until the end of potentially predictable tones. However, these effects could not be replicated in two highly similar follow-up experiments. This questions the validity of the results and partly suggests that the observed effects might be explained in terms of incidental findings. Moreover, the peri-stimulus effects of predictive certainty in Experiment 2 as well as in the collapsed data across all 4 experiments showed a reversed effect around the onset of certainly unpredictable tones. It is therefore not possible to rule out that the observed ERP modulations can be explained by carry-over effects from previous tones which could be driven by alternative mechanisms like different states of refractoriness of underlying neural populations. Furthermore, ERP-modulations in response to omissions of potentially predictable tones, associated with different degrees of predictive reliability, were observed in Experiment 1. These effects could not be replicated in Experiment 2 and 4. As a consequence, at least two major questions need to be answered due to the inconsistencies in the present data. First, are the observed results for ERPs in response to omissions and potentially predictable tones incidental or false positive findings and second, if it can be shown that this is not the case, are the effects really linked to prediction-related parameters and can alternative explanations (like carry-over effects from previous tones) be ruled out using appropriate control conditions?

In order to assess whether the observed findings represent true effects, when planning a replication of the experiments, it should be taken into account that the effects presented in the current thesis are relatively small. Therefore, future investigations should pay special attention to ensure that the recorded data quality is sufficient by following the recommendations provided in Section 5.2.2. Furthermore, measures should be defined to quantify the SNR in order to be able to make informed statements about the data quality and possible consequences for the reliability and validity of the observed effects. For example, it might be helpful to quantify the SNR by calculating the relation between the average magnitude of certain obligatory components, like the P50 and a measure of baseline noise like the root mean squared magnitude of activity in the base-line interval on single-subject level. Moreover, it should be ensured that the number of presented stimuli is appropriate in order to sufficiently cancel spontaneous EEG activity which might interfere with the event-related activity of interest. Introducing additional control conditions as suggested by Bendixen et al. (2015) might further help to rule out possible alternative explanations if expected effects, related to different degrees of predictive certainty, can be regained in future investigations.

Future studies should also aim at systematically investigating pre-stimulus correlates of auditory prediction from different perspectives by utilizing different methods. For example, it might be promising to approach a similar paradigm designed for the use of time-frequency analysis because data from ECoG-measurements suggested that certain prediction-related correlates in the oscillatory activity seem to be induced (Sedley et al., 2016) and hence inaccessible for the event-related potential technique. As introduced in Section 1.2, Wacongne et al. (2012) used an anatomically plausible model of a neural network composed of spiking excitatory and inhibitory neurons. They showed that highly expected stimuli yielded neural preparatory activity shortly before they were expected to occur. Violations of those predictions resulted in a prediction error signal mainly triggered by NMDA-dependent receptor signals. They furthermore demonstrated that ERPs elicited by omissions of highly expected stimuli reflect a "pure" prediction signal which can anatomically be separated from the prediction error response. These signals turned out to be much weaker than signals elicited by prediction error neurons which could further explain the small effects observed in the study by Bendixen et al. (2009) and in the current thesis. Future investigations using procedures like multi-unit and single-unit recordings in the animal model would provide valuable insights about underlying physiological implementations of different aspects of predictive auditory processing which would further guide theories and models about predictive coding. Improved theories and computational models might allow to frame and to test more specific hypotheses about different aspects and different involved signaling cascades of the proposed underlying mechanism.

On a larger scale, the direction of a specific field of research is guided by an aggregate of many different studies investigating the same phenomenon. A single study might report incidental findings but confidence increases and more robust conclusions can be drawn from results if many studies from independent researchers provide results in support of the same theory. There are however certain phenomena associated with the publication process which might systematically bias the overall pattern of published results within a certain field. For example, publication bias is a phenomenon related to the selective publication of studies based on whether results are "positive" or not (Rosenthal, 1979; Iyengar & Greenhouse, 1988). This poses a major problem regarding the validity of published research findings (e.g. Dickersin & Min, 1993; Easterbrook, Gopalan, Berlin, & Matthews, 1991). Analytic procedures like meta-analysis (Glass, 1976) can provide unique insights about the current state of results by aggregating data across many studies within a certain field of research. Moreover, techniques like funnel plots (Light, Pillemer, & Wilkinson, 1984) can indicate whether publication bias is present in the subset of included studies. However, systematic analyses of published results on predictive auditory processing are currently lacking. Such investigations are urgently needed in order to assess the validity of the current state of research. Moreover, authors should be aware of this fact because whether a study is suitable for inclusion in meta-analysis highly depends on the reported measures and statistical parameters. Therefore, authors should always report all statistical results including non-significant results. Moreover, all experimental manipulations and all applied methods should be made transparent. Precise p-values (both, of significant and non-significant results) as well as effect size measures should be reported extensively. If future investigations (e.g. meta-analyses) reveal that the current state of publications on predictive auditory processing is contaminated by publication bias, this might have severe implications for the validity of the published results and the interpretations thereof.

## 5.4 Conclusion

The auditory system is faced with a variety of difficulties which must be overcome in order to extract meaningful information out of the diverse mixture of sound sources which surround us every day. Predictive coding theory provides a good model of how the system might accomplish this difficult task with surprising ease by taking advantage of certain regularities in the sensory input based on which it predicts the upcoming of future sensory signals. The current thesis aimed at characterizing predictive auditory processing by systematically varying different aspects of predictive certainty across multiple levels using event-related potentials within EEG measurements. Graded effects of predictive reliability and predictive accuracy have been observed shortly before the onset of potentially predictable events in two experiments. Furthermore, one experiment revealed effects of predictive accuracy in response to omissions of potentially predictable tones. These results suggest more positive deflections of ERPs associated with higher degrees of predictive accuracy and might indicate that the underlying system engages in predictions in a literal sense and that it flexibly adapts to different degrees of predictive certainty embedded in the sensory context.

However, effects could not be replicated in two follow-up experiments which calls any conclusions that might be drawn from the observed results into question. The inconsistencies of the current results do not permit to draw unequivocal conclusions about predictive auditory processing in audition. Observed ERP-modulations were either very small or expected effects were absent which might at least partially be explained by low statistical power and/or insufficient SNR. Whether these small effect sizes are an indicator of false positive results or whether they reflect some fundamental physiological properties of the proposed mechanism needs to be resolved by future investigations. The current thesis provides some recommendations for future research to improve theoretical considerations and paradigmatic approaches, practical aspects of data acquisition and data analysis, and some general suggestions for the publication of data and the reassessment of published results in order to make informed statements about the validity of reported research findings.

In conclusion, the current thesis improves the original paradigm by Bendixen et al. (2009) in several methodologically important ways. By manipulating different aspects of predictive certainty, the applied paradigm enables to systematically investigate the ecological validity of proposed theories of predictive coding in audition.

This is a crucial step towards the assessment of predictive coding in audition under more realistic conditions. In contrast to previous investigations which commonly contrast fully predictable with certainly unpredictable stimuli, the current thesis provides a systematic variation of predictability across several levels. This parametric manipulation of experimental variables is superior to the former protocol in several aspects: it allows to scrutinize the underlying system under more realistic conditions (since natural signals are rarely either fully predictable or completely unpredictable), it provides means to investigate the dynamics of the underlying system (e.g. whether it behaves linearly or according to an all-or-none principle in response to different degrees of predictive certainty), it might be used to gain information about the direction of effects (e.g. in terms of polarity of electrophysiological results) and it reduces the risk of discovering incidental findings. Future investigations of predictive auditory processing would benefit from adopting this approach since it enables to gain more information about the proposed underlying mechanism and it provides countermeasures against false positive findings at the same time.

# References

Abdallah, S., & Plumbley, M. (2009). Information dynamics: patterns of expectation and surprise in the perception of music. *Connection Science*, *21*(2-3), 89-117.

Acunzo, D. J., MacKenzie, G., & van Rossum, M. C. W. (2012). Systematic biases in early ERP and ERF components as a result of high-pass filtering. *Journal of Neuroscience Methods*, *209*(1), 212-218.

Adams, R. A., Shipp, S., & Friston, K. J. (2013). Predictions not commands: active inference in the motor system. *Brain Structure and Function*, *218*(3), 611-643.

American Electroencephalographic Society. (1994). Guideline thirteen: guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology*, *11*, 111-113.

Anderson, J. E., Aarts, A. A., Anderson, C. J., Attridge, P. R., Attwood, A., Axt, J., ... Barnett-Cowan, M. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251).

Arnal, L. H., & Giraud, A. L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences*, *16*(7), 390-398.

Arnal, L. H., Wyart, V., & Giraud, A. L. (2011). Transitions in neural oscillations reflect prediction errors generated in audiovisual speech. *Nature Neuroscience*, *14*(6), 797-801.

Bakker, M., & Wicherts, J. M. (2011). The (mis)reporting of statistical results in psychology journals. *Behavior Research Methods*, *43*(3), 666–678.

Baldeweg, T. (2006). Repetition effects to sounds: evidence for predictive coding in the auditory system. *Trends in Cognitive Sciences*, *10*(3), 93-94.

Baldeweg, T., Klugman, A., Gruzelier, J., & Hirsch, S. R. (2004). Mismatch negativity potentials and cognitive impairment in schizophrenia. *Schizophrenia Research*, *69*(2-3), 203-217.

Bar, M. (2007). The proactive brain: using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, *11*(7), 280-289.

Barascud, N., Pearce, M. T., Griffiths, T. D., Friston, K. J., & Chait, M. (2016). Brain responses in humans reveal ideal observer-like sensitivity to complex acoustic patterns. *PNAS*, *113*(5), E616-E625.

Barch, D. M., & Yarkoni, T. (2013). Introduction to the special issue on reliability and replication in cognitive and affective neuroscience research. *Cognitive, Affective and Behavioral Neuroscience*, *13*, 687-689.

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012). Canonical microcircuits for predictive coding. *Neuron*, *76*(4), 695-711.

Bauer, M., Stenner, M. P., Friston, K. J., & Dolan, R. J. (2014). Attentional modulation of alpha/beta and gamma oscillations reflect functionally distinct processes. *Journal of Neuroscience*, *34*(48), 16117-16125.

Bekinschtein, T. A., Dehaene, S., Rohaut, B., Tadel, F., Cohen, L., & Naccache, L. (2009). Neural signature of the conscious processing of auditory regularities. *PNAS*, *106*(5), 1672-1677.

Bell, A. J., & Sejnowski, T. J. (1995). An information-maximation approach to blind separation and blind deconvolution. *Neural Computation*, *7*, 1129-1159.

Bendixen, A., Duwe, S., & Reiche, M. (2015). Noise occlusion in discrete tone sequences as a tool towards auditory predictive processing? *Brain Research*, *1626*, 97-107.

Bendixen, A., Prinz, W., Horváth, J., Trujillo-Barreto, N. J., & Schröger, E. (2008). Rapid extraction of auditory feature contingencies. *Neuroimage*, *41*(3), 1111-1119.

Bendixen, A., SanMiguel, I., & Schröger, E. (2012). Early electrophysiological indicators for predictive processing in audition: A review. *International Journal of Psychophysiology*, *83*(2), 120-131.

Bendixen, A., & Schröger, E. (2008). Memory trace formation for abstract auditory features and its consequences in different attentional contexts. *Biological Psychology*, *78*(3), 231-241.

Bendixen, A., Schröger, E., & Winkler, I. (2009). I heard that coming: event-related potential evidence for stimulus-driven prediction in the auditory system. *Journal of Neuroscience*, *29*(26), 8447-8451.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B*, *57*(1), 289–300.

Boutros, N. N., Gjini, K., Eickhoff, S. B., Urbach, H., & Pflieger, M. E. (2013). Mapping repetition suppression of the P50 evoked response to the human cerebral cortex. *Clinical Neurophysiology*, *124*(4), 675-685.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433-436.

Bregman, A. S. (1990). *Auditory scene analysis: the perceptual organisation of sound.* Cambridge, Massachusetts: MIT Press.

Broadbent, D. E. (2013). *Perception and communication.* Oxford: Pergamon Press.

Brodski, A., Paasch, G. F., Helbling, S., & Wibral, M. (2015). The faces of predictive coding. *Journal of Neuroscience*, *35*(24), 8997-9006.

Budd, T. W., Barry, R. J., Gordon, E., Rennie, C., & Michie, P. T. (1998). Decrement of the N1 auditory event-related potential with stimulus repetition: habituation vs. refractoriness. *International Journal of Psychophysiology*, *31*(1), 51-68.

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, *14*(5), 365-376.

Bäss, P., Jacobsen, T., & Schröger, E. (2008). Suppression of the auditory N1 event-related potential component with unpredictable self-initiated tones: Evidence for internal forward models with dynamic stimulation. *International Journal of Psychophysiology*, *70*(2), 137-143.

Callaway, E. (1973). Habituation of averaged evoked potentials in man. In H. V. Peeke & M. J. Herz (Eds.), *Habituation* (Vol. 2, p. 153-174). New York: Academic Press.

Chennu, S., Noreika, V., Gueorguiev, D., Blenkmann, A., Kochen, S., Ibáñez, A., ... Bekinschtein, T. A. (2013). Expectation and attention in hierarchical auditory prediction. *Journal of Neuroscience*, *33*(27), 11194-11205.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(03), 181-204.

Costa-Faidella, J., Baldeweg, T., Grimm, S., & Escera, C. (2011). Interactions between "What" and "When" in the auditory system: temporal predictability enhances repetition suppression. *Journal of Neuroscience*, *31*(50), 18590-18597.

Costa-Faidella, J., Grimm, S., Slabu, L., Díaz-Santaella, F., & Escera, C. (2011). Multiple time scales of adaptation in the auditory system as revealed by human evoked potentials. *Psychophysiology*, *48*(6), 774-783.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *Journal of Neuroscience Methods*, *134*, 9-21.

Deneve, S. (2008). Bayesian spiking neurons I: inference. *Neural Computation*, *20*(1), 91-117.

Deutsch, J. A., & Deutsch, D. (1963). Attention: some theoretical considerations. *Psychological Review*, *70*(1), 80.

Dickersin, K., & Min, Y. (1993). Publication bias: the problem that won't go away. *Annals of the New York Academy of Sciences*, *703*(1), 135-148.

Easterbrook, P. J., Gopalan, R., Berlin, J. A., & Matthews, D. R. (1991). Publication bias in clinical research. *The Lancet*, *337*(8746), 867-872.

Ehrlichman, R. S., Maxwell, C. R., Majumdar, S., & Siegel, S. J. (2008). Deviance-elicited changes in event-related potentials are attenuated by ketamine in mice. *Journal of Cognitive Neuroscience*, *20*(8), 1403-1414.

Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *PNAS*, 7900-7905.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, *39*(2), 175-191.

Fechner, G. T. (1860). *Elemente der psychophysik*. Leipzig: Breitkopf und Härtel.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, *6*(2), 78-84.

Friston, K. J. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1456), 815-836.

Friston, K. J. (2008). Hierarchical models in the brain. *PLoS Computational Biology*, *4*(11), e1000211.

Friston, K. J., & Kiebel, S. (2009). Cortical circuits for perceptual inference. *Neural Networks*, *22*(8), 1093-1104.

Garrido, M. I., Kilner, J. M., Kiebel, S. J., & Friston, K. J. (2009). Dynamic causal modeling of the response to frequency deviants. *Journal of Neurophysiology*, *101*(5), 2620-2631.

Garrido, M. I., Sahani, M., & Dolan, R. J. (2013). Outlier responses reflect sensitivity to statistical structure in the human brain. *PLoS Computational Biology*, *9*(3), e1002999.

Gazzaniga, M., Ivry, R. B., & Mangun, G. R. (2009). *Cognitive neuroscience: The biology of the mind (3rd ed.).* New York: W.W. Norton & Company.

Giard, M. H., Perrin, F., Pernier, J., & Bouchet, P. (1990). Brain generators implicated in the processing of auditory stimulus deviance: a topographic event-related potential study. *Psychophysiology*, *27*(6), 627-640.

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, *5*(10), 3-8.

Grimm, S., Escera, C., Slabu, L., & Costa-Faidella, J. (2011). Electrophysiological evidence for the hierarchical organization of auditory change detection in the human brain. *Psychophysiology*, *48*(3), 377-384.

Haenschel, C., Vernon, D. J., Dwivedi, P., Gruzelier, J. H., & Baldeweg, T. (2005). Event-related brain potential correlates of human auditory sensory memory-trace formation. *Journal of Neuroscience*, *25*(45), 10494-10501.

Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes. *Journal of Cognitive Neuroscience*, *11*(2), 194-205.

Helmholtz, H. v. (1867). *Handbuch der physiologischen optik.* Leipzig: Voss.

Hennighausen, E., Heil, M., & Rösler, F. (1993). A correction method for DC drift artifacts. *Electroencephalography and Clinical Neurophysiology*, *86*(3), 199-204.

Herrmann, C. S., Grigutsch, M., & Busch, N. A. (2005). EEG Oscillations and Wavelet Analysis. In Handy, T C (Ed.), *Event-related potentials: A methods handbook* (p. 229-257). Cambridge: MIT Press.

Herrmann, C. S., Munk, M. H. J., & Engel, A. K. (2004). Cognitive functions of gamma-band activity: memory match and utilization. *Trends in Cognitive Sciences*, *8*(8), 347-355.

Hillyard, S. A., Hink, R. F., Schwent, V. L., & Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science*, *182*(4108), 177-180.

Hodgkin, A. L., & Huxley, A. F. (1945). Resting and action potentials in single nerve fibres. *Journal of Physiology*, *104*(2).

Horváth, J., & Winkler, I. (2004). How the human auditory system treats repetition amongst change. *Neuroscience Letters*, *368*(2), 157-161.

Hughes, H. C., Darcey, T. M., Barkan, H. I., Williamson, P. D., Roberts, D. W., & Aslin, C. H. (2001). Responses of human auditory association cortex to the omission of an expected acoustic event. *NeuroImage*, *13*(6), 1073-1089.

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, *2*(8), e124.

Iyengar, S., & Greenhouse, J. B. (1988). Selection models and the file drawer problem. *Statistical Science*, 109-117.

Izhikevich, E. M. (2003). Simple model of spiking neurons. *IEEE Transactions on Neural Networks*, *14*(6), 1569-1572.

Janata, P. (2001). Brain electrical activity evoked by mental formation of auditory expectations and images. *Brain Topography*, *13*(3), 169-193.

Jensen, O., & Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Frontiers in Human Neuroscience*, *4*, 186.

Jerger, K., Biggins, C., & Fein, G. (1992). P50 suppression is not affected by attentional manipulations. *Biological Psychiatry*, *31*(4), 365-377.

John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science*, epub.

Joutsiniemi, S. L., & Hari, R. (1989). Omissions of auditory stimuli may activate frontal cortex. *European Journal of Neuroscience*, *1*(5), 524-528.

Jung, T. P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., & Sejnowski, T. J. (2000). Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. *Clinical Neurophysiology*, *111*(10), 1745-1758.

Kaiser, J., & Lutzenberger, W. (2005). Human gamma-band activity: a window to cognitive processing. *Neuroreport*, *16*(3), 207-211.

Kanai, R., Komura, Y., Shipp, S., & Friston, K. J. (2015). Cerebral hierarchies: predictive processing, precision and the pulvinar. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*, 20140169.

Kappenman, E. S., & Luck, S. J. (2010). The effects of electrode impedance on data quality and statistical significance in ERP recordings. *Psychophysiology*, *47*(5), 888-904.

Kiebel, S. J., von Kriegstein, K., Daunizeau, J., & Friston, K. J. (2009). Recognizing sequences of sequences. *PLoS Computational Biology*, *5*(e1000464).

Klimesch, W. (1997). EEG-alpha rhythms and memory processes. *International Journal of Psychophysiology*, *26*(1), 319-340.

Klimesch, W., Schimke, H., & Pfurtscheller, G. (1993). Alpha frequency, cognitive load and memory performance. *Brain Topography*, *5*(3), 241-251.

Koelsch, S., Gunter, T., Friederici, A. D., & Schröger, E. (2000). Brain indices of music processing: "nonmusicians" are musical. *Journal of Cognitive Neuroscience*, *12*(3), 520-541.

Kujala, T., Tervaniemi, M., & Schröger, E. (2007). The mismatch negativity in cognitive and clinical neuroscience: theoretical and methodological considerations. *Biological Psychology*, *74*(1), 1-19.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203-205.

Lange, K. (2011). The reduced N1 to self-generated tones: An effect of temporal predictability? *Psychophysiology*, *48*(8), 1088-1095.

Leonard, W. R., Robertson, M. L., Snodgrass, J. J., & Kuzawa, C. W. (2003).

Metabolic correlates of hominid brain evolution. *Comparative Biochemistry and Physiology A-Molecular & Integrative Physiology*, *136*(1), 5-15.

Light, R. J., Pillemer, D. B., & Wilkinson, I. (1984). *Summing Up: The Science of Reviewing Research.* Cambridge: Harvard University Press.

Luck, S. J. (2005). *An introduction to the event-related potential technique.* Cambridge: MIT Press.

Markram, H. (2006). The Blue Brain Project. *Nature Reviews Neuroscience*, *7*(2), 153-160.

Martikainen, M. H., Kaneko, K., & Hari, R. (2005). Suppressed responses to self-triggered sounds in the human auditory cortex. *Cerebral Cortex*, *15*(3), 299-302.

Masicampo, E. J., & Lalande, D. R. (2012). A peculiar prevalence of p values just below .05. *The Quarterly Journal of Experimental Psychology*, *65*(11), 2271-2279.

May, P. J. C., & Tiitinen, H. (2010). Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. *Psychophysiology*, *47*(1), 66-122.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*, 746-748.

Molholm, S., Martinez, A., Ritter, W., Javitt, D. C., & Foxe, J. J. (2005). The neural circuitry of pre-attentive auditory change-detection: an fMRI study of pitch and duration mismatch negativity generators. *Cerebral Cortex*, *15*(5), 545-551.

Mumford, D. (1992). On the computational architecture of the neocortex. *Biological Cybernetics*, *66*(3), 241-251.

Müsseler, J. (2008). Visuelle Wahrnehmung. In J. Müsseler (Ed.), *Allgemeine Psychologie.* Berlin: Springer-Verlag.

Nastase, S., Iacovella, V., & Hasson, U. (2014). Uncertainty in visual and auditory series is coded by modality-general and modality-specific neural systems. *Human Brain Mapping*, *35*(4), 1111-1128.

Neuper, C., & Pfurtscheller, G. (2001). Event-related dynamics of cortical rhythms: frequency-specific features and functional correlates. *International Journal of Psychophysiology*, *43*(1), 41-58.

Nyquist, H. (1928). Certain topics in telegraph transmission theory. *Trans. AIEE*, *47*, 617–644.

Näätänen, R., & Alho, K. (1997). Higher-order processes in auditory-change detection. *Trends in Cognitive Sciences*, *1*(2), 44-45.

Näätänen, R., Gaillard, A. W. K., & Mäntysalo, S. (1978). Early Selective-Attention Effect on Evoked-Potential Reinterpreted. *Acta Psychologica*, *42*, 313-329.

Näätänen, R., Paavilainen, P., & Reinikainen, K. (1989). Do event-related potentials to infrequent decrements in duration of auditory stimuli demonstrate a memory trace in man? *Neuroscience Letters*, *107*(1-3), 347-352.

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, *118*(12), 2544-2590.

Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound - a review and an analysis of the component structure. *Psychophysiology*, *24*(4), 375-425.

Oldfield, R. C. (1971). Assessment and analysis of handedness - Edinburgh Inventory. *Neuropsychologia*, *9*, 97-113.

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*(6), 785-806.

Paavilainen, P., Arajärvi, P., & Takegata, R. (2007). Preattentive detection of nonsalient contingencies between auditory features. *Neuroreport*, *18*(2), 159-163.

Paavilainen, P., Jaramillo, M., Näätänen, R., & Winkler, I. (1999). Neuronal populations in the human brain extracting invariant relationships from acoustic variance. *Neuroscience Letters*, *265*(3), 179-182.

Paavilainen, P., Simola, J., Jaramillo, M., Näätänen, R., & Winkler, I. (2001). Preattentive extraction of abstract feature conjunctions from auditory stimulation as reflected by the mismatch negativity (MMN). *Psychophysiology*, *38*(02), 359-365.

Pannese, A., Herrmann, C. S., & Sussman, E. (2015). Analyzing the auditory scene: Neurophysiologic evidence of a dissociation between detection of regularity and detection of change. *Brain topography*, *28*(3), 411-422.

Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. (1989). Spherical splines for scalp potential and current-density mapping. *Electroencephalography and Clinical Neurophysiology*, *72*(2), 184-187.

Raij, T., McEvoy, L., Mäkelä, J. P., & Hari, R. (1997). Human auditory cortex is activated by omissions of auditory stimuli. *Brain Research*, *745*(1-2), 134-143.

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79-87.

Ritter, W., Vaughan, H. G., & Costa, L. D. (1968). Orienting and habituation to auditory stimuli: a study of short terms changes in average evoked responses. *Electroencephalography and Clinical Neurophysiology*, *25*(6), 550-556.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, *86*(3), 638.

Sabri, M., Liebenthal, E., Waldron, E. J., Medler, D. A., & Binder, J. R. (2006). Attentional modulation in the detection of irrelevant deviance: a simultaneous ERP/fMRI study. *Journal of Cognitive Neuroscience*, *18*(5), 689-700.

Sams, M., Alho, K., & Näätänen, R. (1984). Short-term habituation and dishabituation of the Mismatch Negativity of the ERP. *Psychophysiology*, *21*(4), 434-441.

SanMiguel, I., Widmann, A., Bendixen, A., Trujillo-Barreto, N., & Schröger, E. (2013). Hearing silences: human auditory processing relies on preactivation of sound-specific brain activity patterns. *The Journal of Neuroscience*, *33*(20), 8633-8639.

Schadow, J., Lenz, D., Dettler, N., Fründ, I., & Herrmann, C. S. (2009). Early gamma-band responses reflect anticipatory top-down modulation in the auditory cortex. *Neuroimage*, *47*(2), 651-658.

Schröger, E. (2007). Mismatch negativity: A microphone into auditory memory. *Journal of Psychophysiology*, *21*, 138-146.

Schröger, E., & Wolff, C. (1996). Mismatch response of the human brain to changes in sound location. *Neuroreport*, *7*(18), 3005-3008.

Schröger, E., & Wolff, C. (1998a). Attentional orienting and reorienting is indicated by human event-related brain potentials. *Neuroreport*, *9*(15), 3355-3358.

Schröger, E., & Wolff, C. (1998b). Behavioral and electrophysiological effects of task-irrelevant sound change: A new distraction paradigm. *Cognitive Brain Research*, *7*(1), 71-87.

Schwartze, M., Farrugia, N., & Kotz, S. A. (2013). Dissociation of formal and temporal predictability in early auditory evoked potentials. *Neuropsychologia*, *51*(2), 320-325.

Schwartze, M., Tavano, A., Schröger, E., & Kotz, S. A. (2012). Temporal aspects of prediction in audition: Cortical and subcortical neural mechanisms. *International Journal of Psychophysiology*, *83*(2, SI), 200-207.

Sedley, W., Gander, P. E., Kumar, S., Kovach, C. K., Oya, H., Kawasaki, H., ... Griffiths, T. D. (2016). Neural signatures of perceptual inference. *eLife*, *5*, e11476.

Sengupta, B., & Stemmler, M. B. (2014). Power consumption during neuronal computation. *Proceedings of the IEEE*, *102*(5), 738-750.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 0956797611417632.

Sinkkonen, J. (1999). Information and resource allocation. In R. Baddeley, P. Hancock, & P. Foldiak (Eds.), *Information theory and the brain* (p. 241-254). Cambridge: University Press.

Sokolov, E. N. (1963). Higher nervous functions: The orienting reflex. *Annual Review of Physiology*, *25*(1), 545-580.

Srinivasan, M. V., Laughlin, S. B., & Dubs, A. (1982). Predictive coding: a fresh view of inhibition in the retina. *Proceedings of the Royal Society of London B: Biological Sciences*, *216*(1205), 427-459.

Sussman, E., & Winkler, I. (2001). Dynamic sensory updating in the auditory system. *Cognitive Brain Research*, *12*(3), 431-439.

Swets, J. A. (Ed.). (1964). *Signal detection and recognition by human observers.* New York: Wiley.

Tallgren, P., Vanhatalo, S., Kaila, K., & Voipio, J. (2005). Evaluation of commercially available electrodes and gels for recording of slow EEG potentials. *Clinical Neurophysiology*, *116*(4), 799-806.

Tallon-Baudry, C., Bertrand, O., & Fischer, C. (2001). Oscillatory synchrony between human extrastriate areas during visual short-term memory maintenance. *Journal of Neuroscience*, *21*(20), 177.

Tanner, W. P., & Birdsall, T. G. (1958). Definitions of $d'$ and $\eta$ as Psychophysical Measures. *The Journal of the Acoustical Society of America*, *30*(10), 922-

928.

Tanner, W. P., & Swets, J. A. (1954). A decision-making theory of visual detection. *Psychological Review*, *61*(6), 401-409.

Tikhonravov, D., Neuvonen, T., Pertovaara, A., Savioja, K., Ruusuvirta, T., Näätänen, R., & Carlson, S. (2010). Dose-related effects of memantine on a mismatch negativity-like response in anesthetized rats. *Neuroscience*, *167*(4), 1175-1182.

Tikhonravov, D., Neuvonen, T., Pertovaara, A., Savioja, K., Ruusuvirta, T., Näätänen, R., & Carlson, S. (2008). Effects of an NMDA-receptor antagonist MK-801 on an MMN-like response recorded in anesthetized rats. *Brain Research*, *1203*, 97-102.

Tobia, M. J., Iacovella, V., Davis, B., & Hasson, U. (2012). Neural systems mediating recognition of changes in statistical regularities. *Neuroimage*, *63*(3), 1730-1742.

Tobia, M. J., Iacovella, V., & Hasson, U. (2012). Multiple sensitivity profiles to diversity and transition structure in non-stationary input. *Neuroimage*, *60*(2), 991-1005.

Todorovic, A., & de Lange, F. P. (2012). Repetition suppression and expectation suppression are dissociable in time in early auditory evoked fields. *Journal of Neuroscience*, *32*(39), 13389-13395.

Todorovic, A., van Ede, F., Maris, E., & de Lange, F. P. (2011). Prior expectation mediates neural adaptation to repeated sounds in the auditory cortex: An MEG study. *Journal of Neuroscience*, *31*(25), 9118-9123.

Treisman, A. M. (1964). The effect of irrelevant material on the efficiency of selective listening. *The American Journal of Psychology*, *77*(4), 533-546.

Tremblay, P., Baroni, M., & Hasson, U. (2013). Processing of speech and nonspeech sounds in the supratemporal plane: auditory input preference does not predict sensitivity to statistical structure. *Neuroimage*, *66*, 318-332.

Ullsperger, M., & Debener, S. (2010). *Simultaneous EEG and fMRI: recording, analysis, and application.* New York: Oxford University Press.

Vanhatalo, S., Voipio, J., & Kaila, K. (2005). Full-band EEG (FbEEG): an emerging standard in electroencephalography. *Clinical Neurophysiology*, *116*(1), 1-8.

VanRullen, R. (2011). Four common conceptual fallacies in mapping the time course of recognition. , *2*, 365.

Vaseghi, S. V. (2013). *Advanced signal processing and digital noise reduction.* Chichester, Stuttgart: John Wiley & Sons, Teubner.

Vul, E., & Pashler, H. (2012). Voodoo and circularity errors. *Neuroimage*, *62*(2), 945-948.

Wacongne, C., Changeux, J. P., & Dehaene, S. (2012). A neuronal model of predictive coding accounting for the mismatch negativity. *Journal of Neuroscience*, *32*(11), 3665-3678.

Wacongne, C., Labyt, E., van Wassenhove, V., Bekinschtein, T., Naccache, L., & Dehaene, S. (2011). Evidence for a hierarchy of predictions and prediction errors in human cortex. *PNAS*, *108*(51), 20754-20759.

Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt II [Laws of organization in perceptual forms II]. *Psychologische Forschung*, *4*, 301-350.

Widmann, A., Gruber, T., Kujala, T., Tervaniemi, M., & Schröger, E. (2007). Binding symbols and sounds: Evidence from event-related oscillatory Gamma-band activity. *Cerebral Cortex*, *17*(11), 2696-2702.

Widmann, A., Schröger, E., & Maess, B. (2015). Digital filter design for electrophysiological data - a practical approach. *Journal of Neuroscience Methods*, *250*, 34-46.

Winkler, I. (2007). Interpreting the mismatch negativity. *Journal of Psychophysiology*, *21*(3-4), 147-163.

Winkler, I., & Czigler, I. (1998). Mismatch negativity: deviance detection or the maintenance of the 'standard'. *Neuroreport*, *9*(17), 3809-3813.

Winkler, I., Karmos, G., & Näätänen, R. (1996). Adaptive modeling of the unattended acoustic environment reflected in the mismatch negativity event-related potential. *Brain Research*, *742*(1), 239-252.

World Medical Association. (2013). World medical association declaration of helsinki: Ethical principles for medical research involving human subjects. *JAMA*, *310*(20), 2191-2194.

Yabe, H., Tervaniemi, M., Reinikainen, K., & Näätänen, R. (1997). Temporal window of integration revealed by MMN to sound omission. *Neuroreport*, *8*(8), 1971—1974.

Yordanova, J., Kolev, V., & Kirov, R. (2012). Brain oscillations and predictive processing. *Frontiers in Psychology*, *3*, 416.

Yordanova, J., Kolev, V., & Polich, J. (2001). P300 and alpha event-related desynchronization (ERD). *Psychophysiology*, *38*(1), 143-152.

Yvert, B., Crouzeix, A., Bertrand, O., Seither-Preisler, A., & Pantev, C. (2001). Multiple supratemporal sources of magnetic and electric auditory evoked middle latency components in humans. *Cerebral Cortex*, *11*(5), 411-423.

Zoefel, B., & Heil, P. (2013). Detection of near-threshold sounds is independent of EEG phase in common frequency bands. *frontiers in Psychology*, *4*, 262.

# List of Abbreviations

# List of Figures

# List of Tables